DOCUMENT RESUME

ED 401 303 TM 025 852

AUTHOR	Enright, Mary K.; Oltman, Philip K.
TITLE	Describing Complex Academic Tasks from Six Graduate Disciplines Using Multidimensional Scaling and Clustering.
INSTITUTION	Educational Testing Service, Princeton, N.J.
SPONS AGENCY	Graduate Record Examinations Board, Princeton, N.J.
REPORT NO	ETS-RR-96-14
PUB DATE	Apr 96
NOTE	73p.
PUB TYPE	Reports - Evaluative/Feasibility (142)
EDRS PRICE	MF01/PC03 Plus Postage.
DESCRIPTORS	*Cluster Analysis; Graduate Students; *Graduate Study; Higher Education; Humanities; *Intellectual Disciplines; *Multidimensional Scaling; Physical Sciences; Problem Solving; Social Sciences; *Test Construction; Test Reliability
IDENTIFIERS	*Problem Finding

ABSTRACT

Multidimensional scaling and cluster analysis were used to describe and categorize tasks from six graduate disciplines including academic psychology, applied psychology, English literature, journalism, physics, and electrical engineering. A sample of task descriptions was constructed through interviews with four or five graduate students from each of these disciplines. A rating instrument was designed to describe task goals and to evaluate whether tasks were well or ill structured with respect to various aspects of problem definition and problem solution. Graduate faculty (three to five per discipline) used the rating instrument to characterize a sample of tasks from their disciplines. The scales were found reasonably reliable and were useful in identifying and describing task cluster and how such clusters varied both within and across disciplines. A cluster of short-term problems that were posed by someone other than the student was found in every field, although the other characteristics of this cluster of tasks varied with discipline. In all disciplines except physics, a cluster of complex tasks emerged that was characterized as having multiple objectives that needed to be satisfied. The cluster of complex tasks found in physics was not described clearly by the scales. Problem-finding was an important task characteristic in the social sciences and humanities, but not in the physical sciences. The relevance of multidimensional scaling and clustering to test design is discussed. An appendix contains examples of task descriptions. (Contains 8 figures, 15 tables, and 18 references.) (Author/SLD)

RESEARCH REPORT

DESCRIBING COMPLEX ACADEMIC TASKS FROM SIX GRADUATE DISCIPLINES USING MULTIDIMENSIONAL SCALING AND CLUSTERING

Mary K. Enright
Philip K. Oltman

Educational Testing Service
Princeton, New Jersey
April 1996

Describing Complex Academic Tasks from Six Graduate Disciplines Using Multidimensional Scaling and Clustering

Mary K. Enright

and

Philip K. Oltman

Acknowledgments

We wish to thank a number of people for their contributions to this study. Judith Norback was instrumental in developing and coordinating the interview procedures. We appreciated Diane Wattay's persistence and tact in recruiting students and faculty, and in conducting most of the student interviews. Test development staff including Kelli Boyles, John Economou, Michael Kaplan, and Raymond Thompson provided useful feedback about the task descriptions and our preliminary scheme for classifying tasks. Walter Emmerich provided valuable advice throughout the course of the study. Finally, we would like to express our appreciation to the faculty and students who participated in this research. This research was supported by the Graduate Record Examinations Board and by the Educational Testing Service Research Allocation.

Abstract

The problem of defining the nature and variety of academic tasks is growing in importance as more complex assessment tasks are introduced in many educational contexts. In this study, multidimensional scaling and cluster analysis were used to describe and categorize tasks from six graduate disciplines including academic psychology, applied psychology, English literature, journalism, physics, and electrical engineering. A sample of task descriptions was constructed through interviews with graduate students from these disciplines. A rating instrument was designed to describe task goals and to evaluate whether the tasks were well or ill structured with respect to various aspects of problem definition and problem solution. Graduate faculty used the rating instrument to characterize a sample of tasks from their discipline.

The scales were found reasonably reliable and were useful in identifying and describing task clusters and how such clusters varied both within and across disciplines. A cluster of short-term problems that were posed by someone other than the student was found in every field although the other characteristics of this cluster of tasks varied with discipline. For example, the short-term tasks in engineering and physics were well-structured, requiring the application of established principles, and having objective standards for judging performance. In contrast, a cluster of short-term tasks in English literature were very ill-structured because different conceptual approaches could be relevant, there were alternative methods for accomplishing the tasks, many possible solutions existed, and the student had to define an issue or question to consider. In all disciplines except physics, a cluster of complex tasks emerged that was characterized as having multiple objectives that needed to be satisfied. The cluster of complex tasks that was found in physics was not described clearly by the scales. Problem-finding was an important task characteristic in the social sciences and humanities but not in the physical sciences. The relevance of multidimensional scaling and clustering to test design is discussed.

Describing Complex Academic Tasks from Six Graduate Disciplines Using Multidimensional Scaling and Clustering

The current study was motivated by interest in two related topics: (a) the use of complex tasks in assessments of academic ability or achievement, and (b) disciplinary differences in typical problems and in the kinds of reasoning required. In recent years more complex tasks have been introduced in state and national assessments as well as in the classroom in response to the perceived negative effects of multiple-choice tests on teaching and learning. As interest grows in using complex tasks in assessment, the issue of how such tasks should be defined increases in importance. In the context of standardized academic assessment, test guidelines and specifications are usually developed by committees of experts. Opinions among experts as to what constitutes an "important" and "complex" task are certain to vary. Therefore the development of the guidelines depends on the processes of rational analysis and the development of consensus. In contrast, empirical techniques for identifying and describing important, complex tasks have had a long history in professional and occupational assessment.

The need for systematic descriptions of academic problem-solving tasks can be seen by reflecting on the ways that assessment tasks have been designed in the past. Designing assessment tasks typically involves three levels of analysis (Norback, Rosenfeld & Wilson, 1990). The first concerns the real-world context in which people encounter problems and attempt to solve them. The second or conceptual level has typically involved an abstraction or generalization about the skills, abilities, or processes which are required to solve important tasks. The third level involves designing assessment tasks that are thought to require the identified skills, abilities, or processes. These steps, coupled with a program of validation research, are sufficient if what is desired is a test of decontextualized, general skills. An alternative viewpoint, however, is that assessment is more valid when aspects of the real-world context are preserved in the assessment context (Frederiksen, 1984). One response to the call for more authentic or contextualized assessment is to preserve many of the characteristics of real-world problems in the assessment tasks. Assessment tasks are designed to simulate or imitate real-life problems, an approach that has been common in occupational assessment. This approach, however, may lead to tasks that are too tightly bound to highly specific contexts and that limit generalizability (Messick, 1994). One solution to this problem would be to identify critical features of the real-world tasks that permit the observation of important qualities and to preserve these features in assessment tasks. However, we lack frameworks that allow discussions of problem characteristics at a more abstract or conceptual level. The need for a conceptual system to characterize problems may not be evident until one is faced with the challenge of trying to describe similarities and differences in problems from a variety of different disciplines. The goal of the current research was to develop and evaluate a conceptual framework for classifying the kinds of problem-solving tasks typically encountered in graduate education that could serve as a basis for

designing and characterizing assessment tasks. An assumption underlying this approach is that problem characteristics are important determinants of the reasoning processes involved in the solution of a problem. Some evidence in support of this assumption is summarized in the next section.

Problem structure and its impact on reasoning

At present, generally accepted systems for describing or classifying problems do not exist. Rather, loose dichotomies (ill-structured vs. well-structured, convergent vs. divergent, verbal vs. quantitative, formal vs. informal) or unsystematic labels (diagnosis, interpretation, criticism) typically serve as descriptors of problem types. However, there are two recent discussions of how an information processing model of problem solving (Newell & Simon, 1972) can be extended to describe the characteristics of different classes of problems (Goel & Pirolli, 1991; Perkins, 1990). According to Newell's and Simon's problem-solving model, any given problem can be analyzed in terms of initial and goal states, operators which can be used to move from one problem state to another, and goal tests or evaluation functions that allow the problem-solver to determine when the goal state has been reached. Furthermore, distinctions are often made between a problem-representation or problem-structuring phase in which the problem-solver formulates an interpretation of the problem situation and accesses relevant knowledge, and the problem-solution phase.

Although much of the early work on problem solving was concerned with the solution of formal, well-structured tasks, problem-space analyses have been extended to account for complex, ill-structured, and verbally complex problems (cf. Voss & Post, 1988). Ill-structured problems have under-specified states, goals, and operations, and are typical of the kinds of problems encountered in everyday life. However, the contrast between "ill-structured" and "well-structured" is too broad to serve as a useful way of classifying problems. The need for a "problem theory," a descriptive system to classify problems, as a complement to "problem-solving theory" has been discussed by Perkins (1990). He notes that such a theory would help us to understand (a) in what sense different kinds of problems are difficult or "problematic," (b) the distribution of different kinds of problems in different environments (such as academic disciplines), and (c) how different kinds of problems require different kinds of skills and abilities. Perkins proposes a system to classify a broad range of problems along multiple dimensions based on a problem space analysis. Perkins suggests that problems might be characterized along dimensions such as the stability of the problem (e.g., whether or not outside forces or events can change the problem as it is being solved), transparency (e.g., to what extent a problem solver has access to all the information needed to solve a problem), and simplicity (e.g., whether there are

single or multiple paths to a solution, how much consensus there is about the suitability of a solution).

The power of a classification system based on a problem space analysis is more explicitly illustrated in the research of Goel and Pirolli (1991). Goel and Pirolli engaged in a program of research to describe the characteristics that distinguished design problems from puzzle-like, non-design problems and to demonstrate how such distinctions in the task environment were related to differences in problem-solution processes. One interesting aspect of this research is that commonalities in design problems across disciplines (architecture, curriculum design, engineering) were described. Goel and Pirolli identified twelve characteristics that distinguished design from non-design problems. This analysis of the characteristics of design problems served as the basis for specific hypotheses about differences in the problem-solving and reasoning processes that would be apparent in the solution of design as opposed to non-design problems. Empirical evidence for these hypotheses was obtained through analysis of verbal protocols from expert designers from different disciplines who were observed solving a simplified design problem in their field over a period of two hours. Hypotheses about the need for extensive problem structuring, distinct problem-solving phases, decomposition of the problem, and specific control strategies received support from the empirical analyses.

The work of Goel and Pirolli is important in the present context for a number of reasons. First, they have demonstrated that distinguishing characteristics of certain types of problems, which are common to a number of disciplines, can be identified. Secondly, they provided empirical evidence of links between critical features of design tasks and specific problem-solving or reasoning processes. Finally, they developed tractable design tasks that could be solved within a relatively short period of time (two hours). For our purposes, one limitation of this work is that only one class of problems was analyzed. Another is that there is no consideration of the relationship between discipline characteristics and problem characteristics. Are some types of problems more typical in some disciplines than others? How do problem characteristics, other than type, affect reasoning processes, and are these differences systematically related to disciplinary differences? An accumulating body of research, briefly discussed below, is devoted to understanding the epistemology of various disciplines and the impact of disciplinary characteristics and training on reasoning.

<u>Disciplinary differences and their impact on reasoning</u>

Analysis of disciplinary differences in various epistemological characteristics is a topic of growing interest. Biglan (1973) had college faculty members rate the similarity of disciplines at their colleges and found that disciplines

could be distinguished along three dimensions. These dimensions included (a) the existence of strong paradigms, (b) concern with applications, and (c) concern with life systems. More recently, Donald (1983, 1990, 1991, 1993) has engaged in an ambitious program of research that seeks to understand disciplines with respect to characteristics such as knowledge structure, learning tasks, validation processes, and truth criteria. For example, in one study Donald (1990) interviewed faculty members about the validation processes, truth criteria, and other validation factors characteristic of their disciplines. With respect to validation processes, Donald reported that the use of empirical evidence was emphasized in the natural and social sciences but not in the humanities. In contrast, peer review was a more important validation process in the humanities. Furthermore, faculty members in pure fields were more likely than those in applied fields to use conflicting evidence in validating their work.

This work is particularly relevant in view of other research that documents how the structure and content of a domain are related to the processes involved in problem-solving. Glaser, Schauble, Raghavan, and Zeitz (1992) observed student learning in three different computer-based discovery environments which embodied laws and regularities from the domains of micro economics (law of supply and demand) and physics (electrical circuits, light refraction). The students' task was to induce the laws or regularities governing each environment. The nature of these regularities varies with the domain. Correlational regularities are characteristic of micro economics while functional rules are characteristic of the two domains on physics. Glaser et al. documented differences in evidence-generation activities, interpretative activities, and the use of mathematical and algebraic heuristics that were associated with the nature of the laws governing each domain. This research complements the research of Goel and Pirolli described previously in that it identifies how one characteristic of a specific type of problem (discovery or inquiry) varies with domains and influences the kinds of problem-solving and reasoning processes that are used in problem solution.

The Research Problem and Methodology

The previous discussion suggests that one way to contextualize reasoning is to create assessment problems that embody critical features of the problems typical of a discipline. Problem characteristics are an important determinant of problem-solving and reasoning processes. Critical problem characteristics are likely to vary among disciplines in relation to disciplinary differences in content, structure, and epistemology. However, no systematic or comprehensive description of problem categories and characteristics presently exists. A more abstract description of problems would contribute to the development of assessment problems that better match tasks that are performed in graduate education. It would also provide a basis for grouping disciplines in terms of

critical problem characteristics so that assessment problems appropriate for broad areas rather than specific disciplines can be developed.

The goal of this research was to develop a broad overview of problem-solving tasks typically encountered in graduate education and a framework for classifying these tasks and their attributes. Such a framework would be useful for developing assessment tasks and could be developed through a number of methods. For example, one method for developing task frameworks in assessment relies primarily on the use of expert committees who through discussion arrive at a consensus about the kinds of tasks or the characteristics of tasks that should be included on the assessment instrument. This method links the framework only indirectly, through expert opinion, to an analysis of criterion tasks. Another method is job analysis. A typical procedure in job analysis is to interview one group of experts to construct a list of the knowledge, skills, and abilities (KSAs) thought to be necessary for competent performance and to survey another group to confirm the importance of these KSAs. These procedures are well suited to the development of assessments which measure many discrete skills. However, in the current study we sought to develop a framework that could be used to classify a sample of complex criterion tasks and that could support the development of more complex assessment tasks by delineating some of the critical features of the criterion tasks that would need to be preserved in the assessment context.

Methods for developing classification systems are not well established in psychology though Fleishman and Quaintance (1984) have catalogued issues that need to be considered. As discussed below, some of the issues we confronted in this study included:

1. deciding what constitutes a task and collecting and describing a sample of tasks.
2. identifying a conceptual basis for the classification system
3. specifying relevant task attributes given the conceptual basis
4. deciding on the structural characteristics of the classificatory system
5. evaluating the adequacy of the classification system.

Method

A broad overview of typical problem-solving tasks in graduate education was obtained through interviews with graduate students from six different disciplines. Samples of problem-solving tasks, collected from these students, provided a basis for developing a framework and rating scales for classifying the tasks. Graduate faculty used the rating scales to characterize the tasks collected from students in their discipline. Multidimensional scaling of these faculty ratings was then used to cluster and describe tasks within disciplines.

Disciplines were selected so as to maximize diversity in the sample of problem-solving tasks. The disciplines included in the study were applied psychology, academic psychology, engineering, English literature, journalism, and physics. These disciplines represented the humanities, social sciences, and physical sciences as well as academic and applied disciplines.

We interviewed a small sample (four or five students, three to five faculty members) within each discipline. First, graduate faculty who were willing to participate in the study were identified and they in turn suggested graduate students within their programs who might be interested in participating. After describing participating faculty and students, the three major components of the study, (a) the collection of problem-solving tasks from graduate students, (b) the development of a preliminary classification system for academic tasks, and (c) the application and evaluation of the classification system will be discussed in more detail.

Participating Graduate Faculty and Students

Because the sample of faculty and students we planned to interview was small, we developed a set of criteria for participation directed toward obtaining a diverse sample so that different points of view would be represented. We sought to identify graduate faculty within each field who had some of the following characteristics:

1. a strong interest in graduate education
2. a reputation as a particularly effective mentor
3. an interest in recruiting and supporting graduate students from minority groups
4. a position in a department that is considered as one of the best places to do graduate work in a particular discipline at present
5. a junior faculty member
6. a member of a minority group.

Initially, lists of potential participants were based on suggestions by individuals involved in the development of large-scale subject area tests for the disciplines or by the chairpersons of large or well-known graduate programs. These individuals were then contacted by phone and asked to participate or to recommend colleagues having some of the characteristics listed above. A total of 30 graduate faculty members, five in each of the six disciplines, agreed to participate but five of these did not complete the study. No two graduate faculty members were from the same program or department. Although it was not always possible to enlist faculty members so as to insure that each of the above criteria

was met for each discipline, the faculty sample was reasonably diverse. Some characteristics of the faculty who participated in the study are presented in Table 1.

Graduate faculty were asked to identify graduate students who might be interested in participating in this study. A total of 29 students agreed to participate in the study. Characteristics of the students who participated are described also in Table 1.

Student Interviews

Two interviews were conducted with each student. The purpose of the first interview, which usually took an hour, was to identify examples of tasks that the student had completed in the course of graduate study. Fleishman and Quaintance (1984) note that the notion of a "task" is not well defined in the literature. Our concept of task was broad--a set of activities performed to accomplish a goal--rather than narrow, and assumed that a task has some objective as well as subjective aspects. Furthermore we were concerned with tasks that had an externalized component such as a product or performance that could be evaluated by others. Thus reading a book did not qualify as a task but writing a critique of a book would. Finally, our emphasis was on tasks that were completed in the first two to three years of enrollment in the program so that differences between tasks in masters and doctoral programs would not confound program characteristics with stage of education characteristics.

We developed a task sampling framework to collect a comparable sample of tasks across students and disciplines to the extent possible. During our initial conversations with graduate faculty members we had discussed both explicit and implicit requirements students were expected to fulfill in their graduate program. This information was used to structure our initial interviews with the students. Students were asked to describe tasks they had carried out in the following contexts:

1. course work including lectures, seminars, and laboratory courses
2. teaching responsibilities
3. research
4. degree examinations
5. internships and practicums
6. preprofessional and professional activities.

With each student we identified 5 to 10 tasks for which they could send us sample materials and that would be discussed in more detail during the second interview. Examples of the materials collected include problems or questions from examinations, abstracts or sections from papers or reports, and notes for oral

presentations. In the second interview, the student was asked to describe what was involved in completing each task with respect to the following considerations:

1. Resources - what information sources were used (e.g. textbooks, journals, manuals, professors, other students, own knowledge).
2. Activities and steps involved in carrying out the task.
3. Definition and structure of the task - (e.g., how specific were directions for carrying out task, who defined problem to be solved, how was task completion determined).
4. Duration of task.

On the basis of these interviews a sample of 25 task descriptions was developed for each discipline. Tasks were selected for inclusion in the sample based both on the quality of the information available about the task as well as an attempt to include examples of major classes of tasks carried out in various contexts within each discipline. Examples of task descriptions from each discipline are included in Appendix A.

The Development of the Classification Framework

A preliminary classification system and rating scales drawing upon problem-solving theory, a previous attempt to classify academic tasks (Bloom, 1954), and an examination of the tasks that had been collected were developed by project staff. The rating scales and the sample of tasks were reviewed by test development staff with expertise in the disciplines of interest and the scales were modified in accordance with their suggestions. Graduate faculty were asked to use the scales to rate the sample of tasks from their discipline and to comment on the scales' clarity and relevance to their discipline.

Preliminary Outcomes

The Sample of Problem-Solving Tasks

Most of the tasks in this sample (67%) were carried out in the context of course requirements, reflecting the study's emphasis on the first few years of graduate school when students often complete required courses. Another 11% were tasks completed as a part of independent research or as research assistants. The remaining tasks were distributed among the contexts of teaching (9%), degree examinations (7%), internships (4%), and professional activities (5%).

The most common kinds of tasks described involved answering questions or problems posed by a teacher (37%) and writing long papers or reports (21%).

Also included in the sample were short papers (13%), oral presentations (14%), proposals or plans (7%), and a variety of miscellaneous tasks (5%).

One contrast among disciplines that was very salient in our interviews was the use of problems in electrical engineering and physics as opposed to other disciplines. According to the students we interviewed, much of their time in the first year or two of graduate school is spent solving sets of problems for course homework or examinations. This is reflected in the higher incidence in our sample of tasks involving answering questions or problems posed by a teacher for electrical engineering (52%) and physics (60%) than for the other disciplines (20 to 32%).

This broad categorization of tasks provides an overview of the kinds of tasks included in the sample but obscures the diversity that was present in the sample of tasks. For example the category "short papers or reports" included book reviews, reviews of articles, reflections on readings or responses to presentations, and psychological evaluations of clients. In order to better analyze the diversity in problem characteristics that exists in this sample of tasks, we developed a preliminary classification system for these tasks and rating scales embodying the system.

Classification System for Academic Tasks

In order to determine the distinguishing features of different kinds of tasks we developed a preliminary classification system and an associated rating instrument (see Figure 1). Rather than ask faculty to categorize these complex tasks into discrete categories, the scales were constructed so as to allow multidimensional scaling and clustering of the tasks.

A problem-solving framework provided the underlying rationale for many of the characteristics on the rating scales. Problem-solving theory was used to suggest global differences among tasks, however, and not as a system to carry out a fine-grained comparison of problems in terms of possible representations and solution paths or of the states and operations involved in problem solution (cf. Perkins, 1990).

The classification system embedded in the rating instrument has three components. These include (a) general task requirements, (b) how well-structured different phases of the problem-solving process are, and (c) some contextual factors. The major characteristics on the rating scales correspond to classes of general task requirements including, for example, analysis, inquiry, planning, diagnosis. Although some of these goals have connotations in common with Bloom's (1954) cognitive objectives (e.g. analysis, application) they differ in that

they do not represent a task demand for the use of certain cognitive processes but rather they specify something about the nature of the problem solution the students are expected to produce. These task requirements are more akin to general problem goals. The cognitive processes by which they are achieved would be expected to vary with the expertise of the problem-solver. Because the tasks are complex, more then one requirement can be characteristic of a task.

Other items on the rating scales were intended to probe to what degree various aspects of the problem solution process, including problem-finding, problem-representation, and problem solution, were well-structured or ill-structured. A problem is ill-structured when there are many open constraints that the problem-solver must resolve in the course of problem-solution (Voss and Post, 1988). Academic tasks may vary with respect to whether it is open for a problem-solver to decide (a) what the problem is; (b) what principles, systems, theories, or perspectives might be applied to the problem; (c) which of many alternative solution paths should be taken; or (d) which of many alternative solutions may be best. A task would be well-structured in any of these aspects to the degree that range of alternatives available is limited or constrained by factors such as task instructions or the level of development of knowledge in the discipline. Scale items such as "For the most part, the task is posed or defined by someone other than the student" and "Finding an important issue, topic, or question to consider would be a challenging component of the task for the student" were included to determine the extent to which the student had to find or define the problem. Other characteristics were related to problem representation--"A number of different conceptual systems or approaches might be relevant to the task," and problem solution phases--"Once a conceptual formulation of the task is achieved, the solution is straight-forward or routine," or "There are many different possible solutions for the task."

Some other aspects of the task environment, such as the variety of resources available, and the need to interact with other individuals, were also included. Finally the faculty were asked to rate how similar a task was to those students would encounter in future professional careers, and how informative performance on a task would be about a student's professional development and potential.

Application and Evaluation of the Rating Scales

Analysis

Interjudge agreement. The twenty-five graduate faculty described in Table 1 used the scales to rate the sample of tasks collected from students in their own discipline. To assess the interjudge agreement for the rating data for each of the

six disciplines, Cronbach's alpha was computed among the judges for each task, across the rating scales. Thus, for each discipline, 25 alphas were computed, one for each task. The size of the coefficients indicate the extent to which the set of judges agreed among themselves on the relative magnitude of the ratings of applicability of the scales to each of the 25 tasks. Table 2 displays the 25th, 50th, and 75th percentiles of the distribution of the 25 alphas for each discipline. The magnitude of the coefficients suggests a reasonable amount of agreement among the judges.

Multidimensional scaling and cluster analysis. Non-metric multidimensional scaling solutions for one through four dimensions were performed using the SPSS multidimensional scaling procedure (Norušis, 1993). The s-stress indexes (Takane, Young, & de Leeuw, 1977) for each solution in the data from each discipline are graphed in Figure 2. Although the "elbows" in the curves were not particularly pronounced, we decided to select the two-dimensional solutions for further exploration.

One way to explore the meaning of the dimensions for each discipline is to examine the correlations across the 25 tasks of each of the scales with each of the dimensions. These are displayed in Tables 3 through 8. The most positive and most negative scale correlations with each dimension give a general idea of the way the set of tasks defines the space. However, we gave more attention to the nature of tasks that appeared to group together in the space, in the following way: The multidimensional scaling procedures yielded weights for each of the 25 tasks on the two dimensions. Cluster analysis based on the dimension weights allowed us to explore the two-dimensional task space for "neighborhoods" in which academic tasks might cluster together because they resembled each other in their patterns of ratings by faculty judges. The weights were subjected to cluster analysis using Ward's method (Ward, 1963). The number of clusters to be interpreted for each discipline was determined based upon inspection of the cluster dendrograms for evidence of clear separation of clusters. The six disciplines yielded from 2 to 4 clusters. The 25 academic tasks for each discipline are plotted in two-dimensional space in Figures 3 through 8. The clusters that appeared to stand out clearly are indicated in the figures by polygons enclosing tasks belonging to a given cluster.

Brief précis of each of the task descriptions are shown in Tables 9 through 14, arranged according to the clusters that emerged. To clarify further the nature of the task clusters, each cluster is preceded by a listing of those rating scales that showed mean ratings of 3.0 or greater on a scale ranging from 0 (not appropriate) to 4 (very appropriate).

Results

For all the disciplines the first dimension reflected a contrast between complex, ill-structured tasks requiring inquiry, systemization, planning, interactions with colleagues, and the identification of an important question to consider by the student with tasks where the problem was posed by others and drew on the students' current knowledge. The nature of the second dimension varied more across disciplines. For academic psychology (Table 3), the second dimension opposed tasks that required application of knowledge and implementation of plans with those that required evaluation of ideas while for applied psychology (Table 4) the second dimension reflected a distinction between applied tasks involving a particular case, interaction with people other than colleagues, and diagnosis, with more academic pursuits. The second dimension for English literature (Table 5) reflected a contrast between tasks involving administration and planning and those requiring the formulation of a claim or thesis and analysis and application. In journalism (Table 6) the second dimension contrasted focus on a particular case rather than general issues and availability of alternative methods for accomplishing the task with tasks that required analysis and application of established principles. The second dimension in physics (Table 7) contrasted tasks that were components of larger tasks, focused on the particular, required interactions with others, design, and diagnosis with tasks requiring consolidation and analysis. Finally, the second dimension in electrical engineering (Table 8) contrasted applied tasks that were concerned with the particular rather than the general and involved implementation and administration as opposed to more academic tasks that involved consolidation and formulating hypotheses. The meaning of these dimensions within and across disciplines becomes clearer when the clusters of specific tasks are described below.

Academic Psychology. In academic psychology, three task clusters were identified (Table 9). Cluster 1 primarily consisted of short-term course-related assignments, and examination questions. The major characteristics of this cluster were that the task was defined by someone other than the student, and that the tasks required consolidation and analysis. The tasks in Cluster 2 included research proposals, reports and presentations. These tasks were complex in that they were characterized by multiple general task requirements including consolidation, analysis, application, inquiry, and planning. The challenge of identifying an important issue for consideration, and formulating a claim, confronted the students. In addition alternative conceptual systems and approaches might be relevant to the tasks and these tasks often involved interactions with other professionals and colleagues. Overall, this cluster of tasks requires the student to take responsibility for defining and structuring problems. The tasks in Cluster 3 shared some similarities with the tasks in Cluster 2. Cluster 3 involved research papers that were more focused on summarizing existing knowledge than on

developing new knowledge. These tasks were less complex than the tasks in Cluster 2 in terms of the number of general task requirements that were important task components but they also required student involvement in defining and structuring the problem.

Applied Psychology. Two of the three clusters that were found for applied psychology (Clusters 1 and 2) were very similar to those found for academic psychology (Table 10). Cluster 1 included examination questions, critiques and reviews of books or articles, and preparing a lecture. The major requirements of these tasks included consolidation and analysis and typically the task was defined by someone other than the student. Cluster 2 included tasks such as research papers, proposals, and developing a school-based program. This cluster had much in common with Cluster 2 for academic psychology. The tasks were highly complex in terms of multiple general requirements and there were a number of indications that tasks required structuring by the students. Students needed to define a question and formulate a claim, and a number of alternative systems, methods, and solutions were applicable to the task. Finally, Cluster 3 for applied psychology, which included a variety of tasks such as psychological evaluations, research papers and designing a program, differed from Cluster 3 for academic psychology. This third cluster for applied psychology required inquiry and systemization as well as consolidation and analysis but the problems were posed by others and the task focused on a particular case or example rather than on general issues.

English Literature. The results of the cluster analysis for English literature were particularly interesting (Table 11). More tasks clusters (four) were found for this discipline than for any of the others and more of the scales were rated as highly appropriate descriptors of at least one task cluster than for other disciplines. Clusters 1 and 2 consisted of examination questions and other course related assignments including papers. Cluster 3 involved a few tasks that differed from tasks in the other clusters in that they were primarily planning tasks and involved interactions with colleagues. Finally, Cluster 4 included highly complex tasks such as papers and presentations for courses and seminars. One striking result was that many of the characteristics of the most complex tasks, Cluster 4 (papers and presentations), also were characteristic of Cluster 1 (examination questions). Both Clusters 1 and 4 included tasks that were complex in terms of general task requirements, and unstructured because the students had to define the issue, formulate a claim, and because alternative conceptual systems, methods, and possible solutions were relevant to the tasks. However, these two clusters differed in that problems in Cluster 1 were more likely to be posed by another and students had to rely primarily on their current knowledge to accomplish the task. Additionally, Cluster 4 was characterized by a requirement for systemization--the

elaboration or construction of a system or structure within which information can be interpreted or explained.

Journalism. Three task clusters were found for journalism (Table 12). Cluster 1 consisted of course assignment and examination questions. Only two characteristics were rated as highly appropriate descriptions of these tasks, (a) that the task was defined by someone other than the student and (b) that the student had to rely on current knowledge to accomplish the task rather than consulting a wide variety of other sources. Research-related tasks such as developing survey questionnaires or a presentation on a master's thesis composed Cluster 2. These tasks were more complex than those in the two other clusters, requiring consolidation, application, inquiry, systemization, and the formulation of a claim or thesis. Cluster 3 primarily included news stories, and course papers. The only characteristic that was rated as a very appropriate description of this cluster was consolidation or summarizing, integrating and organizing information about a specific area of knowledge.

Physics. Three task clusters were found for physics (Table 13). However, only 4 of the 24 scales were deemed to be highly appropriate descriptors of the tasks in these clusters. Clusters 1 and 2 consisted primarily of homework and exam problems. These problems were characterized as being posed or defined by someone other than the student and requiring the application of established principles, methods etc. to the solution of the problem. Cluster 1 was also characterized by the existence of explicit and objective standards for judging the quality of performance. Cluster 3 was composed of a more heterogeneous collection of tasks such as writing papers and research proposals, or developing homework assignments for a class. The most salient characteristic of this task cluster was consolidation or the summarization, organization and integration of information about an area of knowledge.

Electrical Engineering. Cluster 1 for electrical engineering was composed of a diverse set of tasks related to writing and presenting papers, laboratory work and teaching laboratory classes (Table 14). This set of tasks was complex in terms of the number of general problem requirements that were used to characterize the tasks. Consolidation, analysis, application, inquiry and systemization were all rated as highly appropriate task characteristics. Two other important descriptors were that a number of different conceptual systems or approaches were relevant to the task and that the tasks often requires interaction with colleagues or students. Cluster 2 consisted mostly of homework assignments and examination problems. This cluster represented well-structured problems that required analysis in terms of and application of established principles, were posed by someone other than the student, had straight-forward or routine solutions, and had explicit and objective

standards for judging performance. This cluster of tasks was similar to the first cluster found in physics.

Task authenticity and predictiveness. Faculty also rated each task with respect to whether the task was very similar to the kinds of tasks students would encounter in their professional careers and whether performance on the task would be highly informative about a student's professional development and potential. The mean ratings on these two scales for the task clusters identified for each discipline are shown in Table 15. Overall, clusters of complex tasks tended to be rated higher on both these scales then clusters of short-term assignments. Furthermore, task clusters tended to be rated slightly higher in terms of similarity than in terms of informativeness. A number of faculty noted in interviews that these scales were difficult to apply because of the wide variety of career paths students might follow.

Comparisons among Disciplines. Some questions that arise in evaluating the use of the rating scales by faculty from different disciplines is whether the scales were equally relevant for all disciplines and the comparability of task clusters across disciplines. Answering these kinds of questions is complicated by the fact that the groups of faculty raters were very small and composed of different individuals for each discipline. Therefore, mean ratings on various scales may confound differences in the relevance of the scales with individual differences in criteria for applying the scales or in tendencies to use extreme values. However, if extreme values are used by the group of faculty within a discipline for at least some of the scales, it seems safe to infer that the ratings reflect judgments of scale relevance and not just differences in criteria. For example, only 4 scales were found to be highly appropriate descriptors (mean >=3) of the tasks in the physics clusters while 16 scales were deemed highly appropriate for describing tasks in English Literature. The fact that the mean ratings for at least some of the scales within clusters were as high for physics (3.7-3.9) as for English Literature (3.7-3.8), suggests that this difference does reflect the relevance of the scales to physics and not just differences in criteria for applying the scales. The rating scales seemed to be inadequate particularly for describing more complex tasks in physics.

Across disciplines, some scales were important in characterizing differences in task clusters in all disciplines while others were important in some disciplines but not others and the differences are related to broad major fields (physical sciences, social sciences, humanities). The scale "For the most part, the task is posed or defined by someone other than the student," was useful in distinguishing clusters of short-term, course related assignments in all disciplines. In physics and electrical engineering, these clusters were characterized also as requiring application of known principles and having explicit and objective standards for judging performance. The latter characteristic did not help differentiate task clusters in the

social sciences and humanities. In academic and applied psychology, the clusters of short-term assignments were similar to each other in that they required consolidation and analysis. The parallelism in the description of the short term tasks between physics and electrical engineering and between academic and applied psychology is worth noting. Different samples of raters from allied disciplines looking at different samples of tasks generated highly similar descriptions of the tasks that differed in meaningful ways from those generated for nonallied disciplines.

Clusters of more complex tasks were characterized by multiple general task requirements in all disciplines but physics. The fact that the cluster of more complex tasks in physics was poorly described by the scales could reflect either the relevance of the scales to this discipline or the adequacy of the task sample. The scales “Finding an important issue, topic, or question to consider..,” and formulating “a claim, thesis, or hypothesis,” also were useful in describing such clusters in the social sciences and humanities but not in the physical sciences indicating that “problem-finding” is a more critical skill in some fields than others.

Discussion

The central purpose of the current study was to explore a task-centered approach as an initial step to inform the test design process. This approach involved an examination of the criterion domain of tasks in which graduate students actually engage. We collected a sample of tasks carried out in a variety of contexts from graduate students in different disciplines. The contexts included course work, research, teaching, degree examinations, internships, and professional activities. To counterbalance the tendency of such an approach to lead to a description of tasks that is too tightly bound to the original context and content, we developed a conceptual framework and rating scales to identify some of the critical features of these "real-world" tasks that might be important to preserve in assessment tasks. Our ideas about critical task features were influenced by recent work on problem-solving theory as well as by Bloom's taxonomy of cognitive objectives. One important characteristic of the rating scales is that they were developed to be applied across disciplines. Another is that they allowed us to evaluate the reliability and usefulness of the scales and the framework in different disciplines. Our expectation was that this framework could be used to group tasks into categories, and to compare the characteristics of different categories of tasks both within and across disciplines.

We found that the scales were reasonably reliable and were useful in identifying and describing task clusters and how such clusters varied both within and across disciplines. A cluster of short-term problems that were posed by someone other than the student was found in every field. However, the other

characteristics of this cluster of tasks varied with discipline. For example, the short-term tasks in engineering and physics were well-structured requiring the application of established principles, and having objective standards for judging performance. In contrast, a cluster of short-term tasks in English literature were very ill-structured because different conceptual approaches could be relevant, there were alternative methods for accomplishing the tasks, many possible solutions existed, and the student had to define an issue or question to consider. In all disciplines except physics, a cluster of complex tasks emerged that was characterized as having multiple objectives that needed to be satisfied. However, the cluster of complex tasks that was found in physics was not described clearly by the scales. Problem-finding was an important task characteristic in the social sciences and humanities but not in the physical sciences. Finally, task clusters in allied academic and applied disciplines were similar to each other but different from those in non-allied disciplines.

In addition to describing the nature of academic tasks in six disciplines, this research has important implications for test development. The research illustrates an alternative approach to developing and evaluating test frameworks and specifications, one that is particularly appropriate for assessments that include complex tasks that vary on a variety of characteristics. For example, test development committees might examine a sample of tasks from their discipline and identify critical features of the tasks. The usefulness and significance of these features for describing both criterion tasks and assessment tasks might be evaluated by having a larger sample of domain experts classify tasks using these features. Such procedures would help clarify and define more precisely what is meant by "complex" or "performance" tasks, establish how meaningful these definitions are, and suggest models for assessment tasks. The differences and similarities that exist between criterion tasks and assessment tasks could be specifically documented and the categories of tasks that would be appropriate for an assessment could be identified. Finally, this approach would allow tasks to be classified on multiple dimensions, the fit of different tasks into the desired categories to be assessed, and the comparability of tasks across different versions of an assessment better documented.

References

Biglan, A. (1973). The characteristics of subject matter in different academic areas. Journal of Applied Psychology, 57, 195-203.

Bloom, B. S. (1954). Taxonomy of educational objectives. Handbook 1: Cognitive domain. New York: Longman, Green

Donald, J. G. (1983). Knowledge structures: Methods for exploring course content. Journal of Higher Education, 54, 31-41.

Donald, J. G. (1990). University professors' views of knowledge and validation processes. Journal of Educational Psychology, 82, 242-249.

Donald, J. G. (1991). The learning task in engineering courses: A study of professors' perceptions of the learning process in six selected courses. European Journal of Engineering Education, 16, 181-192.

Donald, J. G. (1993). Professors' and students' conceptualizations of the learning task in physics courses. Journal of Research in Science Teaching, 30, 905-918.

Fleishman, E. A., & Quaintance, M. K. (1984). Taxonomies of human performance: The description of human tasks. Orlando, FL: Academic Press.

Frederiksen, N. (1984). The real test bias: Influences of testing on teaching and learning. American Psychologist, 39, 193-202.

Glaser, R., Schauble, L., Raghavan, K., & Zeitz, C. (1992). Scientific reasoning across different domains. In E. De Corte, M. Linn, H. Mandl, & L. Verschaffel (Eds.), Computer-based learning environments and problems solving (NATO ASI Series F: Computer and Systems Sciences). Berlin: Springer.

Goel, V. & Pirolli, P. (1991). The structure of design problem spaces. (Technical Report No. DPS-3). Berkeley, CA: University of California at Berkeley.

Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessments. Educational Researcher, 23(2), 13-23.

Newell, A., & Simon, H. A. (1972). Human problem solving. Englewood Cliffs: NJ: Prentice-Hall.

Norback, J., Rosenfeld, M., & Wilson, S. (1990). Identifying the literacy requirements of jobs and job literacy analysis: A new methodology. (ETS Project Report). Princeton, NJ: Educational Testing Service.

Norušis, M. J. (1993). SPSS for Windows professional statistics, Release 6.1. Chicago, IL: SPSS, Inc.

Perkins, D. N. (1990). Problem theory. In V. A. Howard (Ed.), Varieties of thinking. New York: Routledge.

Takane, Y., Young, F. W., & de Leeuw, J. (1977). Nonmetric individual differences multidimensional scaling: An alternating least squares method with optimal scaling features. Psychometrika, 42, 7-67.

Voss, J. F., & Post, T. A. (1988). On the solving of ill-structured problems. In M. T. H. Chi, R. Glaser, & M. J. Farr (Eds.), The nature of expertise. Hillsdale, NJ: Erlbaum.

Ward, J. H. (1963). Hierarchical grouping to optimize an objective function. Journal of the American Statistical Association, 58, 236-244.

Table 1

Characteristics of Study Participants

	Faculty Characteristics			
Discipline	Number Participating	Years teaching graduate students	Number of women	Ethnic Background
Academic Psychology	5	3 - 30	3	Black (1) White (4)
Applied Psychology	4	22 - 30+	1	White (4)
Electrical Engineering	3	4 - 12	2	Black (1) White (2)
English Literature	4	7 - 30+	1	White (3) Black (1)
Journalism	5	12 - 30+	1	White (5)
Physics	4	18 - 30+	1	White (4)
	Student Characteristics			
Discipline	Number Participating	Years in Graduate Program (Range)	Number of women	Ethnic Background
Academic Psychology	5	2 - 4	3	White (5)
Applied Psychology	4	2 - 3	3	White (4)
Electrical Engineering	5	1 - 5	2	Asian Amer. (1) White (4)
English Literature	5	1 - 3	4	Black (1) White (4)
Journalism	5	1 - 8	3	White (5)
Physics	5	2 - 6	2	Asian Amer. (2) White (3)

Table 2

Cronbach's Alpha Among Judges in Each Discipline: 25th, 50th, and 75th Percentiles

	Academic Psychology	Applied Psychology	English Literature	Journalism	Physics	Electrical Engineering
25th percentile	0.78	0.68	0.76	0.73	0.77	0.63
50th percentile	0.85	0.74	0.80	0.80	0.83	0.74
75th percentile	0.87	0.80	0.86	0.84	0.89	0.79

Table 3

Academic Psychology: Correlations of Scales with Dimensions

Scale	Correlation with: Task Dimension 1	Task Dimension 2
1 Consolidation	-0.43	-0.11
2 Analysis	-0.27	-0.31
3 Application	-0.04	0.41
4 Inquiry	-0.87	-0.09
5 Systemization	-0.71	-0.09
6 Designs & Plans	-0.70	0.48
7 Diagnosis/Evaluation	-0.00	-0.71
8 Execution/Implementation	0.01	0.81
9 Reflection	-0.26	0.55
10 Administration	-0.70	0.57
11 Highly Complex	-0.77	0.17
12 Component	-0.07	0.51
13 Posed by Others	0.91	0.03
14 Finding Question	-0.93	-0.21
15 Formulate Hypothesis	-0.85	-0.25
16 Different Concepts	-0.76	-0.58
17 Straightforward	0.73	0.34
18 Many Alternatives	-0.73	-0.31
19 Current Knowledge	0.82	0.14
20 Explicit Standards	0.50	0.38
21 Many Solutions	-0.66	-0.39
22 Colleague Interactions	-0.82	0.44
23 Interactions w/Others	-0.75	0.37
24 Particular Case	0.54	0.03

Table 4

Applied Psychology: Correlations of Scales with Dimensions

Scale	Correlation with: Task Dimension 1	Task Dimension 2
1 Consolidation	-0.55	0.49
2 Analysis	-0.54	0.17
3 Application	-0.53	0.35
4 Inquiry	-0.63	-0.21
5 Systemization	-0.85	0.19
6 Designs & Plans	-0.87	0.06
7 Diagnosis/Evaluation	-0.14	-0.56
8 Execution/Implementation	-0.67	-0.45
9 Reflection	-0.46	0.45
10 Administration	-0.82	-0.22
11 Highly Complex	-0.39	0.26
12 Component	-0.17	-0.31
13 Posed by Others	0.72	-0.19
14 Finding Question	-0.63	0.44
15 Formulate Hypothesis	-0.45	0.37
16 Different Concepts	-0.54	0.38
17 Straightforward	-0.06	0.13
18 Many Alternatives	-0.41	0.51
19 Current Knowledge	0.77	-0.13
20 Explicit Standards	0.03	-0.31
21 Many Solutions	-0.37	0.56
22 Colleague Interactions	-0.80	-0.47
23 Interactions w/Others	-0.66	-0.65
24 Particular Case	-0.02	-0.77

Table 5

English Literature: Correlations of Scales with Dimensions

Scale	Correlation with: Task Dimension 1	Task Dimension 2
1 Consolidation	-0.46	-0.41
2 Analysis	-0.26	-0.61
3 Application	-0.19	-0.60
4 Inquiry	-0.22	-0.41
5 Systemization	-0.53	-0.46
6 Designs & Plans	-0.72	0.45
7 Diagnosis/Evaluation	-0.20	-0.31
8 Execution/Implementation	0.16	0.37
9 Reflection	-0.63	-0.15
10 Administration	-0.46	0.78
11 Highly Complex	-0.53	-0.72
12 Component	-0.32	0.38
13 Posed by Others	0.73	0.13
14 Finding Question	-0.57	-0.42
15 Formulate Hypothesis	-0.07	-0.85
16 Different Concepts	-0.59	0.03
17 Straightforward	0.62	0.45
18 Many Alternatives	-0.43	-0.11
19 Current Knowledge	0.73	0.30
20 Explicit Standards	0.28	0.28
21 Many Solutions	-0.44	-0.16
22 Colleague Interactions	-0.75	0.46
23 Interactions w/Others	-0.24	-0.13
24 Particular Case	0.38	0.16

Table 6

Journalism: Correlations of Scales with Dimensions

Scale	Correlation with: Task Dimension 1	Task Dimension 2
1 Consolidation	-0.45	-0.19
2 Analysis	-0.20	0.63
3 Application	-0.48	0.73
4 Inquiry	-0.89	-0.13
5 Systemization	-0.85	0.30
6 Designs & Plans	-0.70	0.47
7 Diagnosis/Evaluation	0.01	0.30
8 Execution/Implementation	-0.54	0.00
9 Reflection	0.02	-0.41
10 Administration	-0.73	0.00
11 Highly Complex	-0.77	0.51
12 Component	-0.30	0.38
13 Posed by Others	0.85	0.24
14 Finding Question	-0.83	-0.25
15 Formulate Hypothesis	-0.61	0.53
16 Different Concepts	-0.52	0.38
17 Straightforward	0.25	0.06
18 Many Alternatives	-0.44	-0.55
19 Current Knowledge	0.84	0.00
20 Explicit Standards	0.25	0.44
21 Many Solutions	-0.61	-0.47
22 Colleague Interactions	-0.85	0.14
23 Interactions w/Others	-0.61	-0.50
24 Particular Case	-0.03	-0.80

Table 7

Physics: Correlations of Scales with Dimensions

Scale	Correlation with: Task Dimension 1	Task Dimension 2
1 Consolidation	-0.90	-0.30
2 Analysis	0.14	-0.36
3 Application	0.84	-0.01
4 Inquiry	-0.71	-0.02
5 Systemization	-0.81	0.02
6 Designs & Plans	-0.74	0.44
7 Diagnosis/Evaluation	-0.46	0.52
8 Execution/Implementation	0.48	0.04
9 Reflection	-0.52	-0.08
10 Administration	-0.26	0.25
11 Highly Complex	-0.61	-0.24
12 Component	-0.65	0.55
13 Posed by Others	0.92	-0.19
14 Finding Question	-0.91	-0.22
15 Formulate Hypothesis	-0.74	-0.15
16 Different Concepts	-0.81	-0.24
17 Straightforward	0.80	0.06
18 Many Alternatives	-0.88	-0.19
19 Current Knowledge	0.76	-0.23
20 Explicit Standards	0.85	0.00
21 Many Solutions	-0.94	-0.01
22 Colleague Interactions	-0.77	0.04
23 Interactions w/Others	-0.42	0.55
24 Particular Case	0.46	0.71

Table 8

Electrical Engineering: Correlations of Scales with Dimensions

Scale	Correlation with: Task Dimension 1	Task Dimension 2
1 Consolidation	0.71	0.51
2 Analysis	0.51	-0.07
3 Application	-0.11	-0.20
4 Inquiry	0.84	0.32
5 Systemization	0.89	0.22
6 Designs & Plans	0.83	-0.18
7 Diagnosis/Evaluation	0.71	-0.09
8 Execution/Implementation	0.35	-0.64
9 Reflection	0.78	0.29
10 Administration	0.74	-0.44
11 Highly Complex	0.82	-0.01
12 Component	0.56	-0.47
13 Posed by Others	-0.84	-0.01
14 Finding Question	0.81	0.43
15 Formulate Hypothesis	0.57	0.45
16 Different Concepts	0.78	0.32
17 Straightforward	-0.81	-0.26
18 Many Alternatives	0.63	-0.13
19 Current Knowledge	-0.84	0.03
20 Explicit Standards	-0.89	-0.19
21 Many Solutions	0.83	-0.38
22 Colleague Interactions	0.87	-0.32
23 Interactions w/Others	0.56	0.10
24 Particular Case	-0.61	-0.50

Table 9

Academic Psychology Task Clusters

Cluster 1 (High [>=3.0] mean scale ratings on: Consolidation, Analysis, Posed by Others)

PS-01. Take-home preliminary examination
PS-04. Wrote questions for a midterm exam as part of teaching responsibilities
PS-07. Homework problems
PS-08. A 3-hour in-class test
PS-09. Homework problems for a course in statistics
PS-10. A specialty examination
PS-12. Brief critique of assigned readings (journal articles) weekly
PS-15. Weekly reading assignments for a course
PS-20. Final exam in class
PS-21. 3 hour in-class examination for a proseminar
PS-22. Prepared a lecture for an undergraduate course
PS-23. Prepared notes for leading a class discussion
PS-24. Review an article for a journal

Cluster 2 (High mean scale ratings on: Consolidation, Analysis, Application, Inquiry, Designs & Plans, Highly Complex, Finding Question, Formulate Hypothesis, Different Concepts, Colleague Interactions)

PS-02. Presented a paper at a professional meeting
PS-03. Conducted research for a master's thesis
PS-11. Paper summarizing independent research
PS-13. Proposal to the university requesting funding
PS-14. Presentation on first-year research at a mini-convention
PS-16. Proposal for a research grant as an assignment for a seminar
PS-18. Presented a paper at a professional organization

Cluster 3 (High mean scale ratings on: Consolidation, Analysis, Finding Question, Formulate Hypothesis, Different Concepts)

PS-05. Paper describing a research question of interest
PS-06. Paper on "Regulation in Parent-Infant Interactions"
PS-17. Term paper on Self Presentation for a seminar
PS-19. Paper on Idiom Comprehension
PS-25. Write 6 short papers (about 5 pages each) for a course

Table 10

<u>Applied Psychology Task Clusters</u>

Cluster 1 (High [>=3.0] mean scale ratings on: Consolidation, Highly Complex, Posed by Others)

AP-01. Write a 1 hour lecture based on text book material
AP-04. Write a methodological critique (8 to 10 pages) of a journal article
AP-05. Question from a take-home midterm for a seminar
AP-07. Write weekly a critique of a journal article (2 - 3 pages)
AP-08. Item from a 2-hour in-class midterm examination for a lecture course
AP-10. A book review for a journal
AP-14. A question from a 2 hour final exam
AP-15. A question from a 2 hour final exam
AP-17. A question from a 2 hour final exam

Cluster 2 (High mean scale ratings on: Consolidation, Analysis, Highly Complex, Finding Question, Formulate Hypothesis, Different Concepts, Many Alternatives, Many Solutions)

AP-02. A case analysis based on a case described in a book
AP-03. A 27 page paper on cultural bias in IQ testing for a lecture course
AP-09. A grant proposal (26 pages) was written for a seminar
AP-18. Design a research project
AP-20. Design an experiment as a project for a lecture course
AP-21. Develop a program to be implemented in a school system
AP-23. Master's thesis

Cluster 3 (High mean scale ratings on: Consolidation, Analysis, Inquiry, Systemization, Highly Complex, Posed by Others)

AP-06. Interview a new client, write an intake report and present it
AP-11. A poster presentation about master's research at a professional meeting
AP-12. Design a program to meet an important educational or psychological need of people
AP-13. Write a psychological evaluation of an individual based on behavior observations
AP-16. Write psychological assessment reports (about 4 pages) for a lecture course
AP-22. Prepare and give a one hour lecture once a week
AP-24. A 22 page paper on gender biases in the diagnosis of personality disorders
AP-25. Evaluating children who had been identified by teachers as in need of special services

Table 11

English Literature Task Clusters

Cluster 1 (High [>=3.0] mean scale ratings on: Consolidation, Analysis, Application, Inquiry, Posed by Others, Finding Question, Formulate Hypothesis, Different Concepts, Many Alternatives, Current Knowledge, Many Solutions)

EL-01. An item from an in-class final examination
EL-06. Essay question from a general examination on British and American literature
EL-07. A 10-minute report, followed by discussion
EL-12. A question from a general examination on British and American literature
EL-13. One of two papers written for a course using very close textual analysis
EL-20. A question from an open-book, take-home examination
EL-21. A question from a four-hour general examination
EL-24. A question from a final examination for a course in Victorian writers
EL-25. For a course on the American novel, the student wrote a 20-page paper

Cluster 2 (High mean scale ratings on: Inquiry, Posed by Others, Different Concepts, Current Knowledge, Many Solutions, Particular Case)

EL-02. An item from a final examination from a course in 20th Century British Literature
EL-09. A question from a 2-hour final examination
EL-11. A short response paper once a week

Cluster 3 (High mean scale ratings on: Consolidation, Designs & Plans, Finding Question, Different Concepts, Many Alternatives, Many Solutions, Colleague Interactions)

EL-03. The student had to prepare a bibliography for a seminar
EL-08. Prepare a list of questions about the nature of poetry for an undergraduate course
EL-23. Write instructions for an essay assignment on "Television and Its Impact on Society"

Cluster 4 (High mean scale ratings on: Consolidation, Analysis, Application, Inquiry, Systemization, Highly Complex, Finding Question, Formulate Hypothesis, Different Concepts, Many Alternatives, Many Solutions)

EL-04. Identify and analyze the relevance of some historical material to Milton's work
EL-05. A formal 15-minute oral presentation on Kipling's novel *Kim,* followed by discussion
EL-10. A paper using historical and textual evidence to support her claim
EL-14. An oral report for presentation in a graduate seminar
EL-15. A "response paper" critiquing an argument presented in a book
EL-16. Develop an outline for a master's essay
EL-17. A 3-page "response paper" summarizing and critiquing an oral presentation
EL-18. A paper for a course on Milton, applying one or more approaches to an issue
EL-19. Prepare an outline of a proposed term paper to be reviewed by the professor
EL-22. A paper written for a course on post-colonial discourse

Table 12

Journalism Task Clusters

Cluster 1 (High [>=3.0] mean scale ratings on: Posed by Others, Current Knowledge)

JN-01. Homework assignment: Select 2 articles and describe them
JN-05. A question from a 2-hour final examination
JN-08. An item from a midterm examination
JN-09. Item from an in-class midterm which required application of ideas from a text
JN-11. An item from an in-class test
JN-12. Read and summarize current articles from print media
JN-13. Item from a 1-hour mid-term examination
JN-14. Statistical analyses for a professor

Cluster 2 (High mean scale ratings on: Consolidation, Application, Inquiry, Systemization, Highly Complex, Formulate Hypothesis)

JN-02. Master's thesis: Categorize newspaper stories using a computer program
JN-04. Master's research: Design a survey questionnaire
JN-10. Research paper on media attention to a sociopolitical issue
JN-18. For a seminar, student wrote a paper in which he developed some research hypotheses
JN-20. Participate in the design, administration, analysis, and interpretation of a survey
JN-21. Design and administer a questionnaire
JN-23. Oral presentation for a seminar on master's research

Cluster 3 (High mean scale ratings on: Consolidation)

JN-03. Personal narrative reflecting on experiences as an intern
JN-06. Paper for a seminar analyzing gender bias
JN-07. Article for a course on feature writing
JN-15. A news story written for a course in reporting
JN-16. Book reports: 1 book per week
JN-17. A story written for a newspaper when the student was an intern
JN-19. Student did an oral presentation for a class
JN-22 An article written by a student for a reporting class
JN-24. Create assignments for students in an undergraduate journalism course
JN-25. Paper concerning media influence on public opinion

Table 13

Physics Task Clusters

Cluster 1 (High [>=3.0] mean scale ratings on: Application, Posed by Others, Explicit Standards)

PH-01. Final in-class exam
PH-02. In-class final exam
PH-05. In-class midterm exam (applying standard equations)
PH-09. In-class midterm exam
PH-11. Take-home, open-book final exam
PH-14. Homework problem (a probability calculation)
PH-15 Written part of the general exam
PH-18. Homework problem

Cluster 2 (High mean scale ratings on: Application, Posed by Others)

PH-03. Take-home final exam
PH-04. Oral part of the general exam
PH-06. A homework problem
PH-10. A set of problems to be used for practice to prepare for the general examination
PH-13. Lab report which described the research conducted during a lab course
PH-17. Solution to homework that was assigned in an undergraduate class
PH-20. Homework problem
PH-22. Homework problem
PH-25. Generate a graph as a result of data collected in the lab

Cluster 3 (High mean scale ratings on: Consolidation)

PH-07. Write code which was needed to monitor and record information
PH-08. Prepare an abstract of the research she planned to conduct
PH-12. Prepare homework to be assigned to a class
PH-16. Presentation to research group on the current literature
PH-19. Review an end-of-run sheet (a diagnostic tool) and use it to trouble-shoot
PH-21. Generate a research proposal and resources listing
PH-23. Term paper
PH-24. Final class paper

Table 14

Electrical Engineering Task Clusters

Cluster 1 (High [>=3.0] mean scale ratings on: Consolidation, Analysis, Application, Inquiry, Systemization, Different Concepts, Colleague Interactions)

EE-01. Completed a paper for an internship and gave an oral presentation
EE-05. Write a manual as part of research
EE-07. Present a paper at a conference
EE-10. Lab work (building a stand for a crystal laser)
EE-11. Paper to write in lieu of the midterm exam
EE-12. Proposal for a term paper for a class
EE-18. The student read a lab manual and prepared to teach a lab
EE-20. Design and implement specifications for a low-noise differential voltage amplifier
EE-22. The student published a paper during her second year of graduate work
EE-23. Read a lab manual and prepared to teach a lab
EE-25. The student was assigned a class presentation

Cluster 2 (High mean scale ratings on: Analysis, Application, Posed by Others, Straightforward, Explicit Standards)

EE-02. Homework problems
EE-03. Take-home exam
EE-04. In-class final exam, which included 1) figure out what formula applies, and 2) solve it
EE-06. Midterm in-class open-book test
EE-08. Oral part of the preliminary exam
EE-09. In-class midterm exam
EE-13. Homework problem (expressing problem in terms of formulas, and solving the formulas)
EE-14. An in-class test
EE-15. Executing software that simulates a circuit, analyzing any problems and fixing them
EE-16. As part of teaching responsibilities, wrote the solutions to class homework problems
EE-17. Questions about an amplifier, applying circuit theory to the specific problem.
EE-19. Homework problems
EE-21. Wrote a structured program for use in teaching a class
EE-24. Written part of the degree examination

Table 15

Task Authenticity and Predictiveness

Discipline	Cluster	Ratings on 0 to 4 Scale: Similar to kinds of tasks encountered in profession	Informative about student's professional potential
Academic Psychology	1	2.6	1.8
	2	3.9	3.4
	3	2.4	2.4
Applied Psychology	1	2.7	2.5
	2	3.3	3.1
	3	3.4	3.1
English Literature	1	2.6	2.3
	2	1.5	1.8
	3	3.8	2.3
	4	3.4	3.0
Journalism	1	2.4	2.0
	2	2.9	3.0
	3	2.7	2.6
Physics	1	2.1	1.9
	2	2.7	2.2
	3	3.1	2.7
Electrical Engineering	1	3.4	3.5
	2	2.2	2.3

Figure Captions

Figure 1. The rating instrument used by graduate faculty to characterize task descriptions collected from graduate students in various disciplines.

Figure 2. S-Stress versus number of dimensions.

Figure 3. Academic psychology: Clusters in 2-dimension space.

Figure 4. Applied psychology: Clusters in 2-dimension space.

Figure 5. English literature: Clusters in 2-dimension space.

Figure 6. Journalism: Clusters in 2-dimension space.

Figure 7. Physics: Clusters in 2-dimension space.

Figure 8. Electrical engineering: Clusters in 2-dimension space.

RATING SCALE Rater No. _____ Task No. _____

To what degree is each of the following an appropriate characterization of the task. ***(Please circle a number on the scale.)***

Major Characteristics	SCALE: Not Appropriate ... Very Appropriate
1. *Consolidation* - an important component of the task is to summarize, organize, or integrate information about, or to reflect on, a specific area of knowledge or topic.	0.....1.....2.....3.....4
2. *Analysis* - task requires the analysis of a situation in terms of established principles, methods, classification systems, critical systems, or theoretical positions.	0.....1.....2.....3.....4
3. *Application* - task requires the application of established principles, methods, classification systems, critical systems, or theoretical positions to the solution of a problem.	0.....1.....2.....3.....4
4. *Inquiry* - an important component of the task is to search for, collect, and evaluate evidence in order to describe, interpret, or explain something.	0.....1.....2.....3.....4
5. *Systemization* - an important aspect of the task is to elaborate or construct a system or structure within which information can be ordered, interpreted, integrated, or explained. . . .	0.....1.....2.....3.....4
6. *Designs & Plans* - an important goal of the task is the development of a design, plan, or set of directions.	0.....1.....2.....3.....4
7. *Diagnosis/Evaluation* - task requires the determination of what is wrong with a product, system, or set of ideas.	0.....1.....2.....3.....4
8. *Execution/Implementation* - task involves the execution or implementation of previously developed plans or carrying out directions.	0.....1.....2.....3.....4
9. *Reflection* - an important component of the task is self-evaluation or reflection about one's role in events or about what has been learned from an experience.	0.....1.....2.....3.....4
10. *Administration* - task requires the coordination of activities of a number of people, and management of resources.	0.....1.....2.....3.....4
Other Characteristics	
11. The task is highly complex.	0.....1.....2.....3.....4
12. The task is a component of a larger task.	0.....1.....2.....3.....4
13. For the most part, the task is posed or defined by someone other than the student.	0.....1.....2.....3.....4
14. Finding an important issue, topic, or question to consider would be a challenging component of the task for the student.	0.....1.....2.....3.....4

Next page, please

Rater No. _____ Task No. _____

Other Characteristics (continued)	SCALE: Not Appropriate ... Very Appropriate
15. An important aspect of the task is that the student needs to formulate a claim, thesis, or hypothesis.	0.....1.....2.....3.....4
16. A number of different conceptual systems or approaches might be relevant to the task.	0.....1.....2.....3.....4
17. Once a conceptual formulation of the task is achieved, the solution is straight-forward or routine.	0.....1.....2.....3.....4
18. There are many alternative methods for accomplishing the task.	0.....1.....2.....3.....4
19. In accomplishing the task, students must rely primarily on their current knowledge rather than consult a wide variety of other information sources.	0.....1.....2.....3.....4
20. Explicit and objective standards exist for judging the quality of performance on the task.	0.....1.....2.....3.....4
21. There are many different possible solutions for the task.	0.....1.....2.....3.....4
22. The task requires interactions with colleagues, other professionals, or students.	0.....1.....2.....3.....4
33. The task requires interactions with people other than colleagues, professionals, or students.	0.....1.....2.....3.....4
24. The task is concerned with a particular instance, case, or example rather than with general issues.	0.....1.....2.....3.....4
25. The task is very similar to the kinds of tasks that students will subsequently encounter in their professional careers.	0.....1.....2.....3.....4
26. Performance on this kind of task is likely to be highly informative about a student's professional development and potential.	0.....1.....2.....3.....4

Ideas for other characterizations of task: ______________________________

____

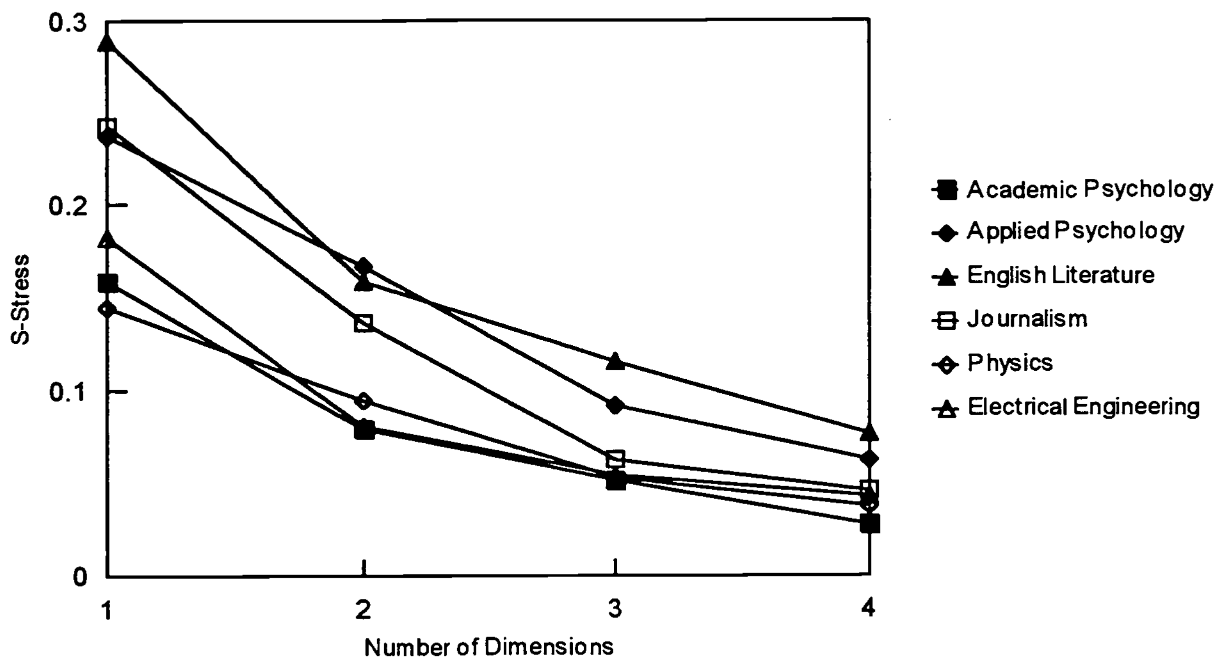
____

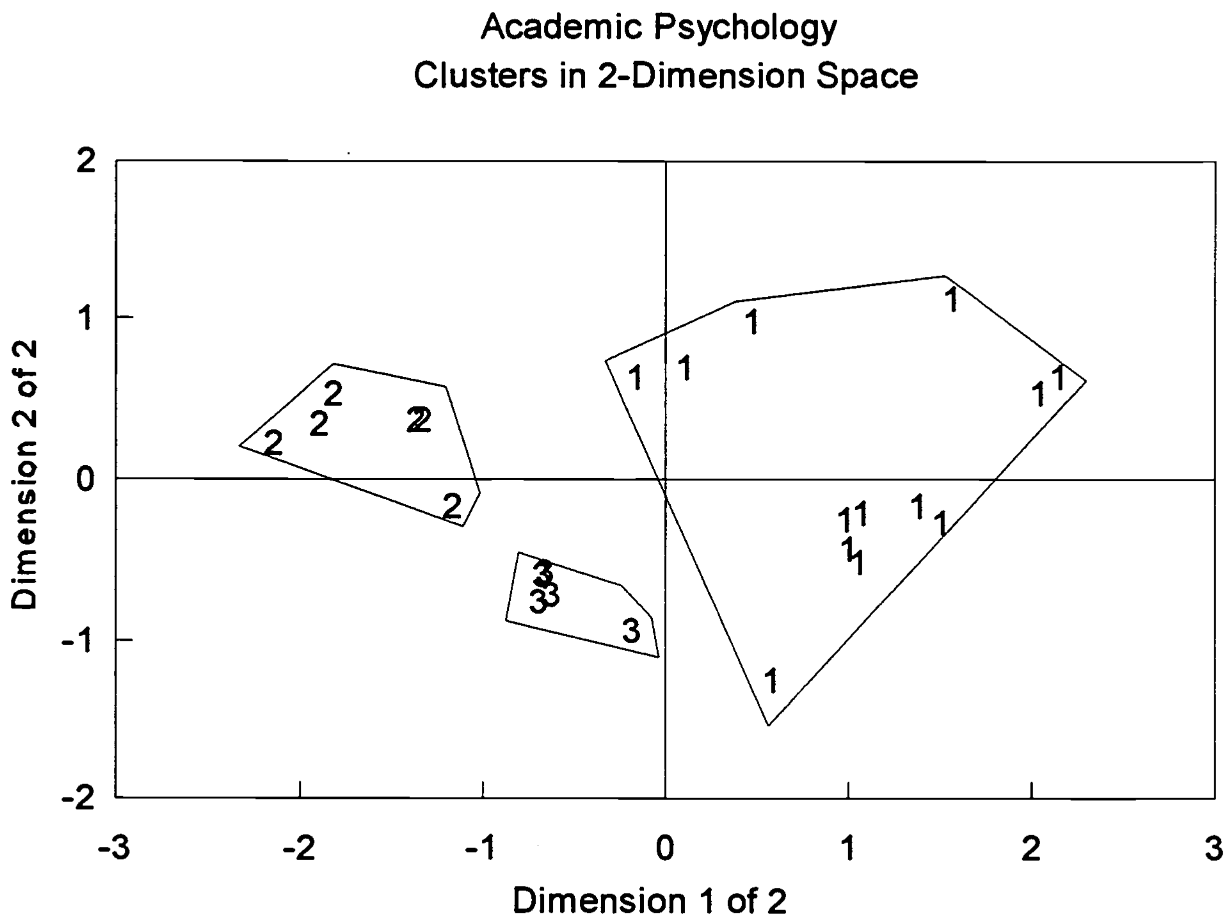
Other comments: ______________________________

____

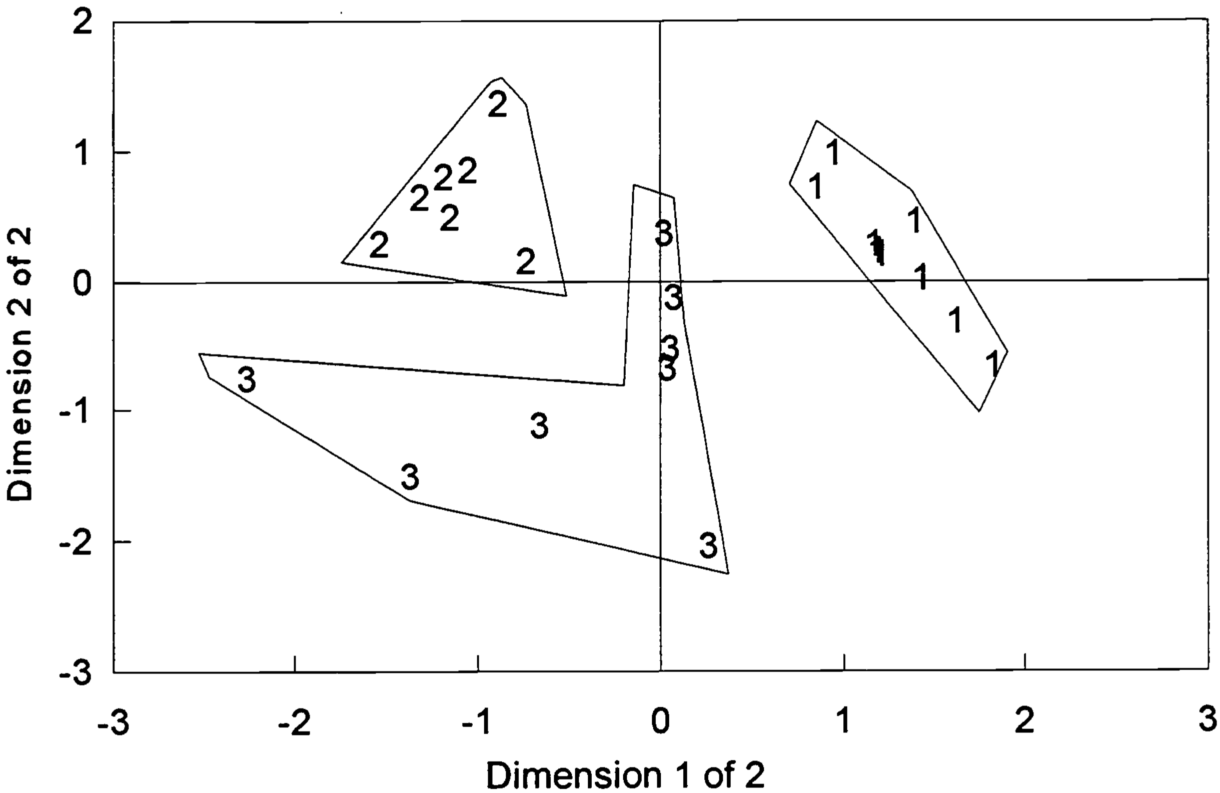
____

S-Stress vs. Number of Dimensions

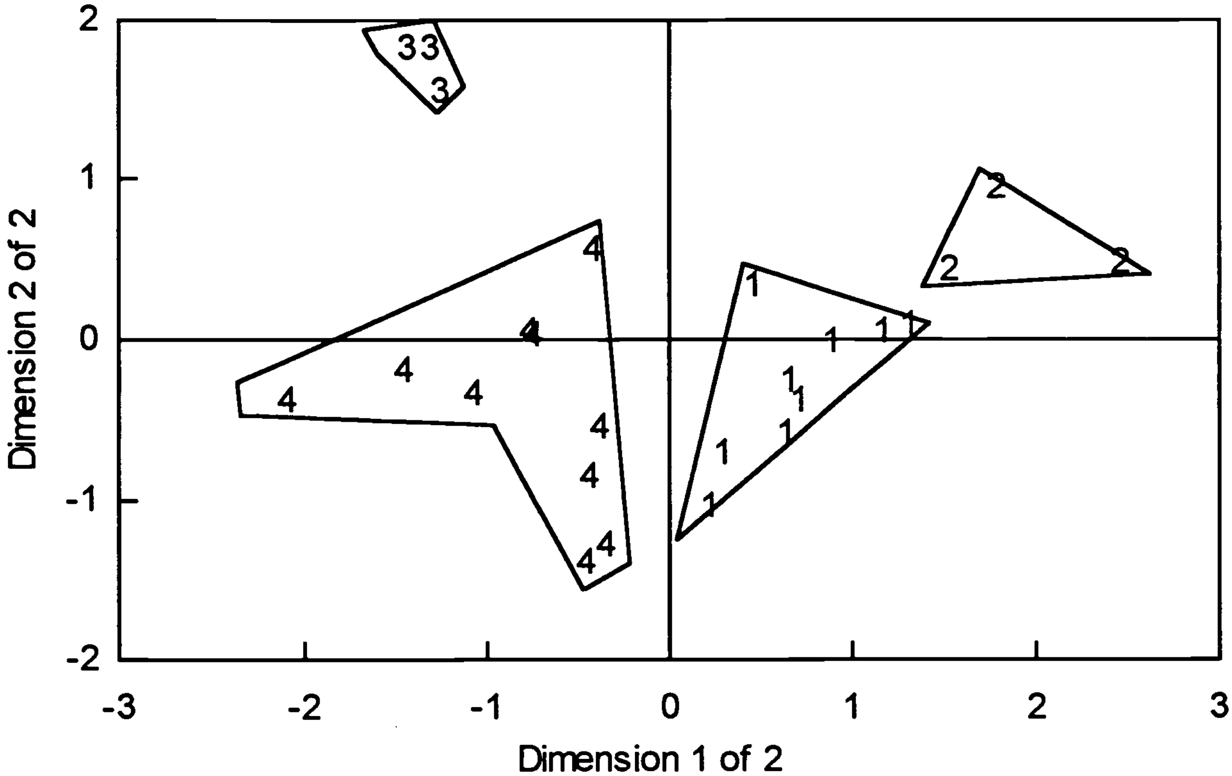


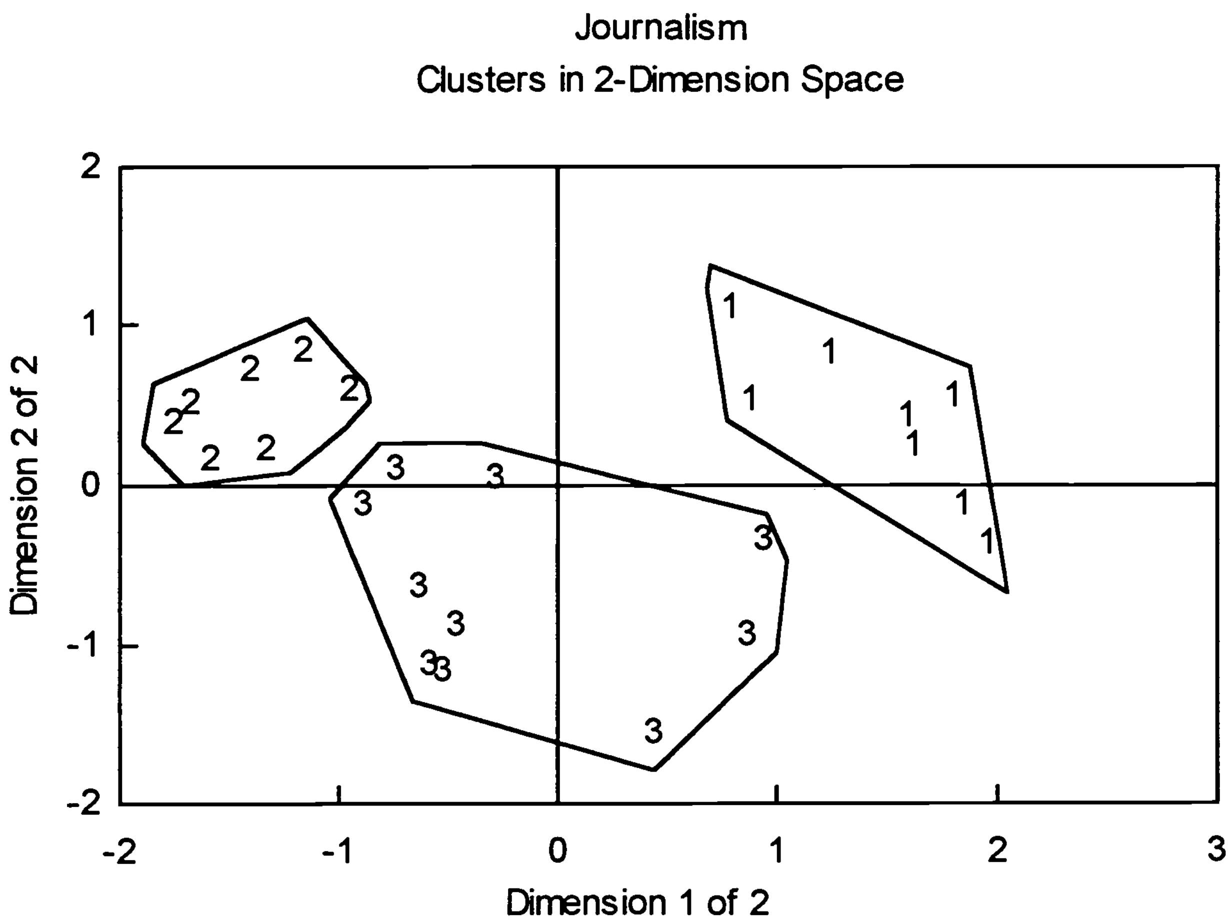

Academic Psychology
Clusters in 2-Dimension Space
Dimension 2 of 2
Dimension 1 of 2
2
1
0
-1
-2
-3
-2
-1
0
1
2
3

Applied Psychology
Clusters in 2-Dimension Space

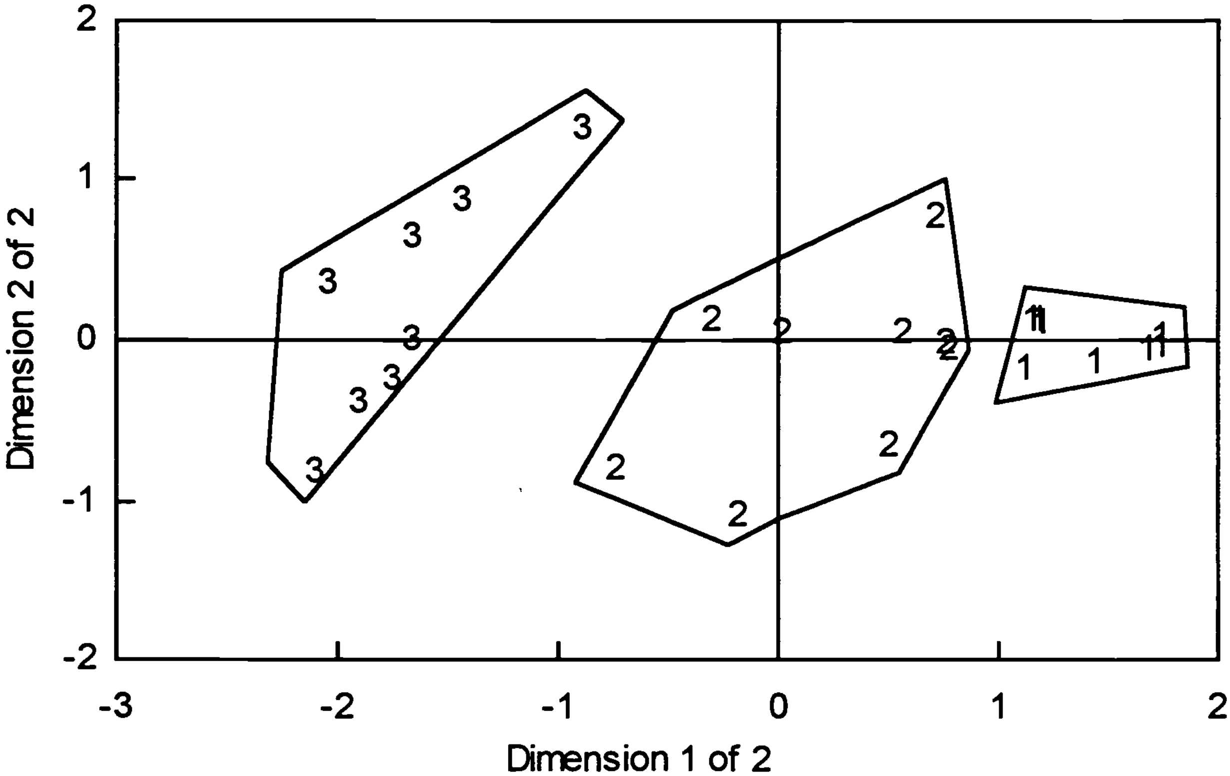


English Literature
Clusters in 2-Dimension Space

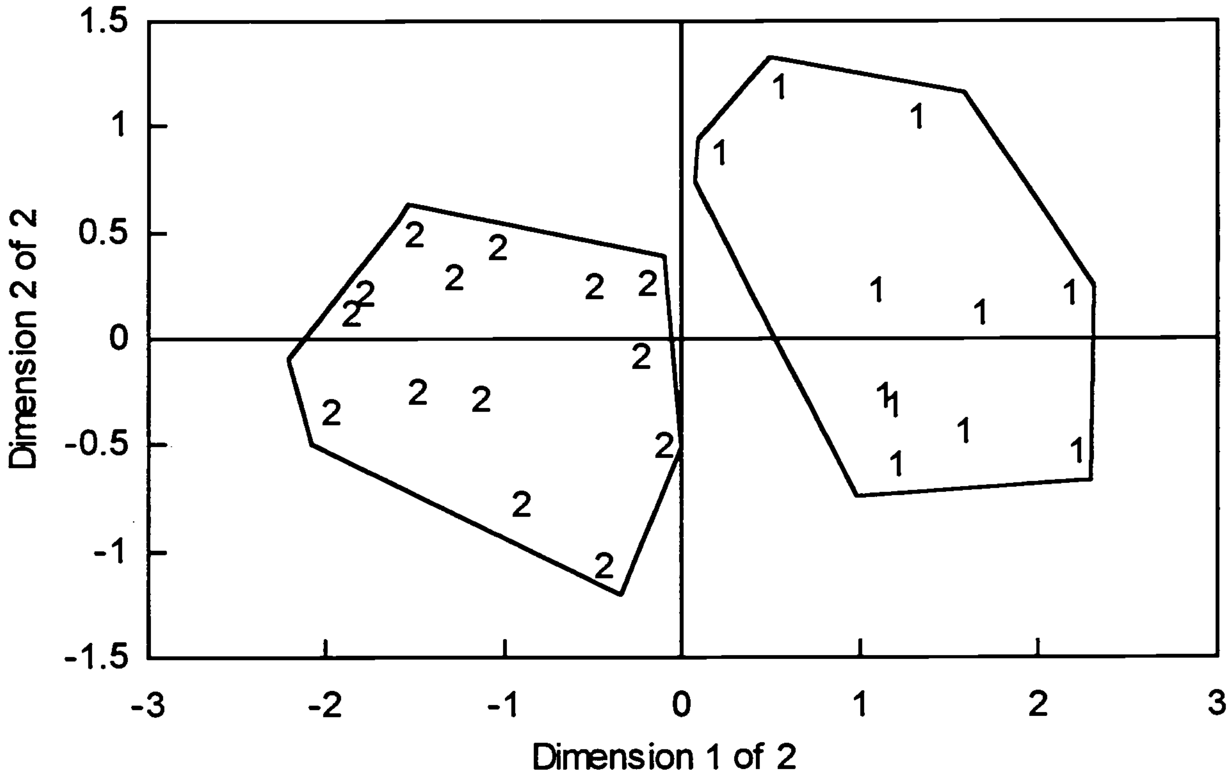



Journalism
Clusters in 2-Dimension Space
Dimension 2 of 2
Dimension 1 of 2
2
1
0
-1
-2
-2
-1
0
1
2
3

Physics
Clusters in 2-Dimension Space



Electrical Engineering
Clusters in 2-Dimension Space



Appendix A

Examples of Task Descriptions from Six Disciplines

AP-06

Task Description

As part of an internship at a testing center and outpatient clinic, the student had to interview a new client, write an intake report and present it to the rest of the staff.

Materials

attached - intake summary

AP-

INTAKE SUMMARY

Patient Name: [redacted]

PRESENTING COMPLAINT: [redacted] is a 33 year old, divorced, black mother of one who was referred by Easter Seals. The patient reports that she is depressed and that this started two years ago when she was diagnosed as having diabetes. At that time she was working at the YWC as a housekeeper but subsequently was forced to quit because of complications with her illness. It was also at this time that she seperated from her husband of several years. The patient states that she is still in love with him but does not want him around because he uses drugs and she does not want her daughter seeing this. Currently she lives with her thirteen year old daughter, who she feels has been hurt by all of this because they can no longer do the things they used to. She also claims that her daughter has had an "attitude" lately and she feels as if they are growing apart and she is losing her, which also upsets her.

RELEVANT HISTORY: The patient reports that at birth her biological mother left her at the hospital and she was subsequently taken in by a foster mother. She claims that her foster mother used to abuse her, both mentally and physically. For example, the patient reported that one time her mother hit her in the head machete and almost took out her eye. She feels that her foster mother would do these things to "make sure she remembered who was boss". At age eleven she found out about what her real mother and contacted her. She claims that they used to meet in the park every Wednesday until one day when she just did not show up. About five years later they had contact again and have remained in touch. However, she states that she is not close to her real mother and that they hardly ever talk. The patient claims that she does not care to keep in contact with her mother or her family because she has three close friends whom she considers her family. However, she states that she does want to know why her mother did not care for her but did for her 2 sisters, 2 half sisters and 5 brothers.

CURRENT MEDICAL STATUS: The patient has diabetes, high blood pressure, sleep apnea, and neuropathy. She takes mendications for these illnesses but is not sure of the names.

MENTAL STATUS: Overall, the patient was easily engaged and made good eye contact. Her thoughts were clear and there was no evidence of glaring cognitive difficulties. Furhtermore, the patients affect was appropriate for the content of the conversation.

DIAGNOSTIC IMPRESSION:
Axis I: Dysthymia
Axis II: Deferred
Axis III: Diabetes, High Blood Pressure, Sleep Apnea, Neuropathy
Axis IV: Moderate
Axis V: present GAF 65 past GAF 70

TREATMENT RECOMMENDATIONS: The patient stated that she would feel more

AP-06

comfortable in individual therapy because she is afraid how other people in a group will respond to her situation. However, the patient stated that if that is the best form of treatment for her she will consider going into a group.

AP-07

Task Description

For a course, the student had to write weekly a critique of a journal article (2 - 3 pages). The student reported that this was an easy task because the articles weren't very difficult.

Materials

attached - 2 page critique

The article "Long-Lasting Alterations in Behavior and Brain Neurochemistry Following Continuous Low-Level LSD Administration" by W. King JR., and G. Ellison, examines whether or not the after-effects of continuous, low level administration of LSD in rats is greater than that of daily injections of the same amount of drug. This study could be of theoretical importance because many of the after-effects of LSD are similar to psychotic reactions of a schizophrenic type. The results suggested that LSD has persisting neurotoxic effects when administered in a continuous low-level fashion. After examining this article closely, I have discovered several issues that need to be considered while reading it.

The first issue that needs to be addressed has to do with the subjects. Although they were 69 albino rats, which makes them more homogeneous than human subjects, there was twice as many male rats represented in the study compared to female rats. Ethically, human's could not have been subjects in this study. Secondly, the subjects were not divided into groups that took cross gender differences into account. The 48 male rats were assigned to only the behavioral tests and the 21 female rats were assigned to only the autoradiography tests. Even though I am not an expert in rat physiology, I can imagine that the same dosage of LSD will effect both sexes in a different manner behavioral and neurologically. Therefore, not to have included both genders in the two different tests limits that tests results to that sex rat and possibly to

that sex in humans (but this is even a bigger generalization).

A second issue that needs to be addressed has to do with the tests themselves. In the open field test, used to assess for average social distance, two rats were placed in a circular enclosure with a flat black interior and a floor divided by white lines into 22 cm squares. The location of each rat was recorded every 12 seconds for 1 minute. In my opinion, after watching the subjects for a while it would be easy to miss a couple of seconds here and there, which in turn could effect the data, especially since this had to be done for twelve rat pairs over a period of ten days. Also, this data could have been easily fudged to make the results come out the way the experimenters wanted them to.

The third issue that needs to be addressed has to do with the fact that base rates were never established for any of the subjects social behavior. Although there was a control group, in an experiment such as this it would have been wise and fairly easy to have established base rates. The reason for this has to do with the fact that the authors were studying the after-effects of a drug. Thus, how can you know the after effects of LSD when you do not know how the rats behaved before the drug was administered.

Lastly, the authors concluded that their findings are congruent with earlier suggestions linking LSD's hallucinogenic effects with alterations in temporal and limbic structures. However, the authors do not say how hallucinations were defined by the behavior in rats. Thus, it is very possible that the alterations in the temporal and limbic structures were due to the LSD itself and not necessarily the drugs hallucinogenic effects.

EE 5

Task Description

The student was asked to write a manual as part of his research. He completed 2 1/2 months of research, and then spent 2 weeks writing the manual. Resources included prior knowledge, internalized knowledge base, the professor, and journal articles. Steps included reviewing previous research in the area, formulating their approach, writing a computer program to use as a modeling tool, and then, after the research was completed, writing the manual.

The student found the assignment medium-to-difficult. No information was provided about how to solve the problem.

Materials

attached (2 pages)

Foreword

The first section of this user's guide provides a description of UALGRL and the prob lems it can analyze. A tutorial that guides the user through an interactive session wit UALGRL is also given in the first section. The second section of this manual is a referenc guide to UALGRL. It provides a description of the input data required by UALGRL an the output data UALGRL can generate. Included in the reference section is a descriptio of the error messages produced by UALGRL and their most probable causes and solu tions. For the advanced user, a discussion of the mathematical models, their limitations and implementation in UALGRL is also supplied in the reference section.

This version of UALGRL features the following improvements:

- Flexible input/output processing.
- User friendly mesh generation.
- Sources and sinks can lie on the edge of the plane.
- The plane can be divided into regions with different values of conductivity.

`Please note that this font is used to indicate input and output from UALGRL.`

EE 13

Task Description

The student was asked to solve a homework problem. Resources included prior knowledge, the textbook, and class notes. The steps in solving it included: 1) formulating the word problem in terms of formulas, and 2) solving the formula.

The student found the task of medium difficulty. It would have been harder if the directions were less specific. The student was told what to use to solve the problem. The scope of the problem did not change as it was solved. There was only one approach to the correct solution to the problem.

Materials

attached (1 page)

1.11† Silicon atoms are added to a piece of gallium arsenide. The silicon can repl
trivalent gallium or pentavalent arsenic atoms. Assume that silicon atoms a
ionized dopant atoms and that 5% of the $10^{10}\ \mathrm{cm}^{-3}$ silicon atoms adde
gallium atoms and 95% replace arsenic atoms. The sample temperature is
(a) Calculate the donor and acceptor concentrations.
(b) Find the electron and hole concentrations and the location of the Fe
(c) Find the conductivity of the gallium arsenide assuming that lattice sca
dominant.
See Table 1.3 for properties of GaAs.

1.12† (Dielectric relaxation in solids.) Consider a homogeneous one-carrier cond
conductivity σ and permittivity ϵ. Imagine a given distribution of the mobi
density $\rho(x,y,z; t = 0)$ in space at $t = 0$. We know the following facts from
magnetism, provided we neglect diffusion current:

$$\nabla \cdot D = \rho; \quad D = \epsilon\mathscr{E}; \quad J = \sigma\mathscr{E}; \quad \nabla \cdot J = \frac{-d\rho}{dt}$$

1A (a) Show from these facts that $\rho(x,y,z; t) = \rho(x,y,z; t = 0)\, e^{-t/(\epsilon/\sigma)}$. This resu
that uncompensated charge cannot remain in a uniform conducting mat
must accumulate at discontinuous surfaces or other places of nonunifor
~~1B~~ (b) Compute the value of the *dielectric relaxation time* ϵ/σ for intrinsic
for silicon doped with 10^{16} donors cm^{-3}; and for thermal SiO_2 with σ =
$(\Omega\text{-cm})^{-1}$.[12]

1.13† Because of their thermal energies, free carriers are continually moving thro
crystal lattice. While the net flow of all carriers across any plane is zero at
equilibrium, it is useful to consider the directed components that balance t
The component values are physically significant in that they measure the
of current that can be delivered by diffusion alone. This would be relevan
example, one were able to unbalance the thermal equilibrium condition by inte
all carriers flowing in a given direction. By considering that $J_x = -qn_0 v_x$, sh
the current in a solid in any random direction resulting from thermal proce

$$J = \frac{-qn_0 v_{th}}{4}$$

where v_{th} is the mean thermal velocity and n_0 is the free-electron density.
(*Hint*. Consider the flux through a solid angle of 2π steradians.)[12]

1.14* Calculate the wavelengths of radiation needed to create hole-electron pairs in i
germanium, silicon, gallium arsenide, and SiO_2. Identify the spectrum ran
infrared, visible, UV, and X ray) for each case.

1.15† The relation between D and μ is given by

$$\frac{D}{\mu} = \frac{1}{q}\frac{dE_f}{d(\ln n)}$$

for a material that may be *degenerate*. (That is, the Fermi-Dirac distribution f
must be used since the Fermi level may enter an allowed energy band.) Sh
this relation reduces to the simpler Einstein relation $D/\mu = kT/q$ if the ma
non-degenerate so that Boltzmann statistics can be used.

EL-07

Task Description

For a graduate seminar, the student prepared a 10 minute report, delivery of which was to be followed by 15 minutes of questions from and discussions among the seminar participants. The report is essentially an article-review: it presents the main argument of an essay in which a critic offers a feminist-structuralist interpretation of Charlotte Lennox's *The Female Quixote*, and conveys some of the textual evidence offered by that critic.

Materials

First paragraph from notes for oral report -

Langbauer begins by noting that for 18th century authors, romance was everything the novel was not: "the contrast between them gave the novel its meaning," as "the utility of romance consisted precisely in its vagueness; it was the chaotic negative space outside the novel" (29). Lennox's Female Quixote "structures its story on the contrast between the novel and romance," so that the "silly extravagances of romance that Arabella illustrates are meant as a foil for the novel's strengths" (29). Lennox thus points to the fictiveness of romance and defines the novel negatively as "real and true." She tries to define The Female Quixote as a stable and controlled novel in part by deriding romance as nonsense. Lennox thus displaces the fictiveness of all fiction onto romance and away from the novel.

EL-13

Task Description

One of two papers written for a course. The student argues that, in Flaubert's *Madame Bovary*, characters and plot are not conveyed through the devices that, in the student's view, were conventionally used in the 19th Century Realist Novel, with the result that the reader must "follow the shifting narrative tone" in order to make sense of these elements. The student develops and supports this point through very close textual analysis of an English translation of Flaubert's work and through references to other critics' readings both of the passages analyzed and of Flaubert's work as a whole.

Materials

First paragraph of paper -

Madame Bovary begins in medias res, with no preamble to identify the narrator, and only the barest indication of setting, thus disorienting the reader and forcing him or her to look to the narrative tone for clarification and understanding. The indeterminate narrative voice moves in and out of the minds of the characters, and through their world with varying focal lengths, at times withdrawing altogether to offer an ironic comment. This indeterminacy requires a willingness on the part of the reader to follow the shifting narrative tone, without expecting the kinds of characterization usually found in realist novels. We understand these characters not through the typology of Balzac or the psychology of George Eliot, but from the way Flaubert's irony treats them. The reader must abandon expectations of a reasonable and readable world in Madame Bovary, and instead, follow the play of language in Flaubert's stylization.

JN-02

Task Description

Master's thesis - Categorized newspaper stories using a computer program and checked its reliability against human coders. The project took the student about 5 months and resulted in a 70-page paper.

The student developed the idea for topic in an earlier course. He then wrote a 20-page proposal including literature review and a description of the methodology and discussed the proposal with advisers. Conducting the research involved collecting a sample of newspaper articles, developing a categorization system, identifying key words in stories to be used as basis for categorization, developing a simple computer program to count occurrences of keywords and to categorize stories. The computer program was used to categorize stories and the student's colleagues also categorized the same stories. A statistical analysis comparing agreement of computer and human categorization was conducted. During this project the student met frequently with two advisers to discuss progress. The report was written up in stages as work progressed.

The most difficult aspect of the project was that the student had a general idea but found it difficult to conceptualize what he wanted to do more specifically.

Materials -

Abstract

Traditional methods of categorizing newspaper stories use human coders. Using a key-words list, analyses can be made by a computer program with high reliability and an absence of bias. The program, called GENCA, analyzed 609 randomly selected news stories from seven newspapers, classifying 95 percent of them into nine categories of news. As a check on reliability, five human coders classified 50 randomly selected stories, agreeing with GENCA an average of 72 percent. Agreement among themselves averaged 75 percent. GENCA agreed with at least one coder on all but six stories and agreed with the majority of coders in 82 percent of the cases.

JN-08

Task Description

An item from a midterm examination. The student found the item relatively difficult because it required analysis using a particular system.

Materials -

Item. For the following hypothesis, provide concept names, a theoretical definition for each concept, an operation definition for each concept, a theoretical linkage for the hypothesis, and an operational linkage for the hypothesis. The hypothesis is:

The more television a person watches, the fewer books he or she reads.

Physics 1

Task Description

The student was given a final in-class exam. Twenty minutes were allowed to respond to each page of problems, one of which is attached. Resources included prior knowledge and internalized knowledge base.

The problem was considered of medium difficulty. The directions were very specific. The scope of the problems did not change as they were solved, and there were not alternative approaches to solution.

Materials

attached (one page)

Page of final exam for Quantum Mechanics I

21

(b) Derive the total angular momentum eigenstate $|\frac{1}{2}\frac{1}{2}\rangle_{jm}$ from the fact that it must be orthogonal to the state you derived in (a). $|\frac{1}{2},\frac{1}{2}\rangle_{jm} = a|0,\frac{1}{2}\rangle + b|1,-\frac{1}{2}\rangle$

$0 = \langle \frac{1}{2},\frac{1}{2} | \frac{3}{2},\frac{1}{2}\rangle = \frac{\sqrt{2}a + b}{\sqrt{3}}$ $\quad a = -b/\sqrt{2}$ $\quad \frac{b^2}{2} + b^2 = 1 \Rightarrow b = \sqrt{\frac{2}{3}}$

$$\therefore \boxed{|\tfrac{1}{2},\tfrac{1}{2}\rangle_{jm} = \frac{-|0,\frac{1}{2}\rangle + \sqrt{2}|1,-\frac{1}{2}\rangle}{\sqrt{3}}}$$

✓

3. Still considering two particles, one of spin 1 and one of spin $\frac{1}{2}$:

(a) Construct in terms of $\mathbf{S}^{(1)}$ and $\mathbf{S}^{(2)}$ the projection operators $P_{3/2}$ and $P_{1/2}$.

$J = S^1 + S^2 \qquad J^2 = (S^1)^2 + (S^2)^2 + 2S^1\cdot S^2$

$S^1\cdot S^2 = \frac{J^2 - (S^1)^2 - (S^2)^2}{2} = \frac{j(j+1) - 1(1+1) - \frac{1}{2}(\frac{1}{2}+1)}{2}$

for $j=\frac{1}{2}$ $\quad S^1\cdot S^2 = -\frac{2}{2} = -1$; $j = 3/2$ $\quad S^1\cdot S^2 = \frac{\frac{3}{2}(\frac{3}{2}+1) - 2 - \frac{3}{4}}{2}$

$\frac{\frac{15}{4} - \frac{3}{4} - \frac{8}{4}}{2} = \frac{1}{2}$

$$\boxed{P_{3/2} = \frac{S^1\cdot S^2 + 1}{3/2}} \qquad \boxed{P_{1/2} = \frac{S^1\cdot S^2 - \frac{1}{2}}{-3/2}}$$

✓

so $P_{3/2}|j=3/2\rangle = 1$
$P_{3/2}|j=1/2\rangle = 0$

so $P_{1/2}|j=1/2\rangle = 1$
$P_{1/2}|j=3/2\rangle = 0$

(b) Use the projection operators you found in 3 to construct directly the state $|\frac{1}{2}\frac{1}{2}\rangle_{jm}$. (If you could not do 3, indicate in general terms how, given such projection operators, you could construct the state.)

$$|\tfrac{1}{2},\tfrac{1}{2}\rangle_{jm} \propto \underbrace{-\frac{2}{3}\left[S^1\cdot S^2 - \frac{1}{2}\right]}_{P_{1/2}} |0,\tfrac{1}{2}\rangle$$

$$\propto -\frac{2}{3}\left[\underbrace{S_x^1 S_x^2 + S_y^1 S_y^2}_{\frac{1}{2}(S_+^1S_-^2 + S_-^1S_+^2)} + \cancel{S_z^1 S_z^2}_{0} - \frac{1}{2}\right]|0,\tfrac{1}{2}\rangle$$

easier $\longrightarrow$

$$-\frac{2}{3}\left[\frac{1}{4}(S_+^1 + S_-^1)(S_-^2) + \frac{1}{4}(S_+^1 - S_-^1)(S_-^2) + \frac{1}{2}|0,\tfrac{1}{2}\rangle\right]$$

$S_x = \frac{1}{2}(S_+ + S_-)$
$S_y = \frac{1}{2i}(S_+ - S_-)$

S_-^2 ... here

$$|\tfrac{1}{2},\tfrac{1}{2}\rangle_{jm} \propto \frac{\sqrt{2}}{2}|1,-\tfrac{1}{2}\rangle - \frac{1}{2}|0,\tfrac{1}{2}\rangle \qquad \left(\begin{array}{l} S_-^2 = 1 \\ S_+^1 = \sqrt{2} \end{array}\right)$$

✓

normalizing to $\langle\frac{1}{2},\frac{1}{2}|\frac{1}{2}\frac{1}{2}\rangle = 1$, we obtain

$$|\tfrac{1}{2},\tfrac{1}{2}\rangle_{jm} = \frac{-|0,\frac{1}{2}\rangle + \sqrt{2}|1,-\frac{1}{2}\rangle}{\sqrt{3}} \quad \text{as before}$$

Physics 21

Task Description

A student acting as a research assistant was given one month to generate a research proposal and resources listing. Resources included prior knowledge, internalized knowledge base, professors, other students, and articles. The steps included were: 1) researching the motivation...why do this experiment? 2) estimating or calculating the counting rates (how long it would take to do the measurement or gather the data), 3) deciding what equipment to use, 4) repeating #2 and #3 above for different types of equipment to optimize results, and 5) writing the proposal, using all this information.

The student found the task of medium difficulty. No information was provided with the assignment; the professor decided the topic, but the student did the rest of the structuring. The process changed as the problem was solved, but the formulas didn't change. There were an infinite number of alternative approaches to solution. A harder problem might have involved trying to measure something that was harder to measure.

Materials

attached (2 pages; resources listing and introduction to proposal)

PROPOSAL INFORMATION

Beam Area: A

Secondary Channel: P^3 West or East *

Beam Requirements:

Type of Particle: π^-
Momentum Range: 192 to 287 MeV/c (98 to 180 MeV)
Momentum Bite: 1 to 4 % ($\Delta p/p$)
Solid Angle: 7 msr
Spot Size: 0.8-cm waists (FWHM)
Emittance: Standard P^3
Intensity: (0.5 to 1.5) x 10^8 π^-/s
Beam Purity: Standard P^3 performance
Targets: Standard A2 target

Primary Beam Requirements: 800 MeV, 0.8-mA average proton beam

Time Required for Experiment

Installation (no beam): one four week period
Collection of Data: 128 hours
Calibration and Normalization: 120 hours
Degrader study: 24 hours
Target manipulation: 24 hours

Total: 296 hours

Scheduling: Summer, 1991

Major LAMPF Apparatus required: Double Focussing magnetic spectrometer used in Exps. 99, 309, 337, 750, 783, 856, 859, 884 and 1026, rare gas handling system for ^{3}He, LAMPF standard computer system and assorted NIM and CAMAC equipment from LEEP.

Shielding and Enclosures Required: Usual concrete shielding at P3 plus existing shielding for the spectrometer.

Special Services Required: 500-kW magnet power supply (2000 A at 250V) regulated to 10^{-4}, 65° F deionized cooling water with Δp = 300 psi at 40 gal/min, cable tray between counting house and spectrometer, liquid nitrogen.

Space Required: Location for spectrometer in P^3 West or East 54 in. downstream of final quad, or as close as possible, with power and cooling water delivered as in previous experiments and with space for ion chamber and profile monitor. P^3 beam must be delivered to its dump without obstruction and the spectrometer must rotate freely.

---

* We prefer P^3 West but only because the spectrometer can be placed closer to the last quadrupole and still rotate from 20° to 130°.

Introduction

Although several experiments [1-5] have observed the isobaric analog transitions induced by pion single charge exchange in ^3He and ^3H, it is unclear what conclusions may be drawn from their results. At this time there is only one measurement at center-of-mass π^0 angles forward of 80°, performed with the π^0 spectrometer,[3] at an incident energy of 200 MeV, and thus there is only one angle-integrated cross section available. The remaining experiments[1,2,4,5] all detected the recoiling nucleus, generally limiting their range of observation to angles greater than 80°. In addition, the recoil measurements at 200 MeV [1] appear to be a factor of two or three times lower than would be expected from a smooth extrapolation of the forward-angle measurement at 200 MeV (see Fig. 1). However, because of the uncertainties in the absolute magnitudes of the reported cross sections, it is difficult to determine how serious the disagreement is. The existing experimental information on both the $^3He(\pi^-,\pi^0)^3H$ and the $^3H(\pi^+,\pi^0)^3He$ reactions is shown in Figure 2, along with cross sections calculated by Gerace *et. al.*[6]

We propose to measure the $^3He(\pi^-,\pi^0)^3H$ differential cross section at incident pion energies of 100, 142, and 180 MeV and in an angular range of roughly 30° to 130° in the π^0 center-of-mass system by detecting the recoil triton. These three energies are chosen to match energies at which reliable elastic scattering measurements have been made.[7-9] In this energy range the existing charge exchange data comprise three rather limited measurements of the charge-symmetric reaction $^3H(\pi^+,\pi^0)^3He$.[2,5] A complete, systematic study of the charge exchange reaction at these energies should challenge theoretical calculation to describe both elastic scattering and single charge exchange at the same time. The measurements will be performed with a magnetic spectrometer, equipped with an array of silicon solid state detectors at its focal plane, viewing a cooled ^3He gas target.

Scientific Motivation

The goal of this experiment is to study the pion-nucleon charge exchange reaction, $\pi^- p \rightarrow \pi^0 n$, with the nucleon bound inside a nucleus. It is commonly accepted that three-body nuclear wave functions are well understood; therefore, the isobaric analog transition between ^3He and ^3H should be well suited to theoretical examination. A variety of methods has been employed to calculate the differential cross section for this transition.[6,10-17] While Glauber multiple scattering calculations [6,10,14] agree reasonably well with existing data above 250 MeV, optical potential calculations better describe the data below 200 MeV.[12,15-17] The poor agreement of the optical potential calculations with

PSY-10

Task Description

The student took a specialty examination. The two-hour exam consisted of 3 questions, all of which had to be completed.

Materials -

One question from the examination.

Social psychologists have emphasized that personality is contructed in social interaction and then, depending on the interactional circumstances, perseverates until further interactions require its reconstruction. Are there no limits, set by a person's personality or other givens, on the personalities that the individual can construct? Suggest what some of these limits might be, where conceptually they come from, and how they can be given a role in social psychological accounts of the social construction process.

PSY-14

Task Description

The student gave a presentation on her first-year research at a mini-convention of faculty and students. With the help of an adviser, the student had identified a research topic and designed a study. She developed a questionnaire, pretested it on a small sample, and then used it on a larger sample. After analyzing the data the student wrote a 25 page paper and prepared for an oral presentation.

Materials -

attached - first page of notes for oral presentation

My first year project explores the relationships between gender, instrumental and expressive self-schemas, and helping behavior.

1. **Primary interests: explore gender & ways people help**

One of my primary interests was to explore whether or not gender i related to ways people help others. In 1986, Alice Eagly and Maureen Crowley published a **meta-analysis** on gender and helping behavior. In the meta-analysis, it becomes clear that most of the research concerning helping behaviors has involved **short-term encounters with stranger** where help is needed (for example, a motorist is stranded because his car has broken down, or an old woman has fallen). In addition, most of the helping research has concentrated on **willingness to help**, but not ways of helping. Eagly and Crowley found that overall, **men were more likely** than women to help but the authors suggest that this finding may be explained by considering the types of helping situations that have been investigated. Typically, many of these short-term encounters involving strangers can be **threatening to the helper**. For example, if a car is broken down on a side street and it is 3am, of course is seems more likely that a man would be willing to pull over and try to help whereas a woman might question her safety and choose to continue driving down the road. **One of my goals** then was to study helping behaviors in situations that were not threatening to the helper.

In 1988, **Eisenberg and her colleagues** examined gender-related traits and helping strangers in non-emergency situations. They found that women were more likely to help strangers than men in a non-emergency situation.

In hopes to expand this area of research on gender and helping, I attempted to explore not only likelihood of helping, but

NOTICE

REPRODUCTION BASIS

☒ This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

☐ This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").